

PT e.g. cytokine and cell proliferation/differentiation activity, the
 PT immune system and hematopoiesis regulating activity -
 PS Disclosure: Page 487-490; 557pp; English.

XX Human cDNA clones represented in AAF98374 - AAF98489 encode secreted
 CC proteins AAB90667 - AAB90750. The cDNA clones are isolated from various
 CC tissue types, and may be used in the prevention, treatment and diagnosis
 CC of diseases associated with inappropriate protein expression. The
 CC polypeptides and nucleic acids may be used as nutrients or to modulate
 CC cytokine and cell proliferation/differentiation activity and may also be
 CC involved in modulation of the immune system. The cDNA sequences
 CC proteins, their agonists and/or antagonists exhibit haematopoiesis
 CC regulating activity; tissue growth activity; activin/inhibin activity;
 CC chemotactic/chemokinetic activity; haemostatic and thrombolytic
 CC activity; receptor/ligand activity; anti-inflammatory activity; and/or
 CC haematopoiesis activity; cadherin/tumour suppressor activity; and/or
 CC tumour inhibition activity. Included in the invention are probes
 CC represented in AAF98490 - AAF98572 which are specific for the cDNA clones
 CC encoding the secreted proteins.

XX Sequence 975 AA:

Query Match 93.6%; Score 5103.5; DB 22; Length 975;
 Best Local Similarity 94.4%; Pred. No. 0;
 Matches 972; Conservative 0; Mismatches 3; Indels 55; Gaps 1;

QY 1 MREALLIVFTLLHFGAGPDESEPTISHGNYTKQYFVGHKQGRNTTQRRLDIOM 60
 DB 1 MREALLIVFTLLHFGAGPDESEPTISHGNYTKQYFVGHKQGRNTTQRRLDIOM 60
 QY 61 IMINGTLYIARADHITVDITDTSHEIYCSKKLTWKSQADVDTCRMKGRKDCQHN 120
 DB 61 IMINGTLYIARADHITVDITDTSHEIYCSKKLTWKSQADVDTCRMKGRKDCQHN 120
 QY 121 IKVLLKNDALFYCGTNAFPCSNRYKMDLFEFGDEFGMARCPDAAHANVALPADG 180
 DB 121 IKVLLKNDALFYCGTNAFPCSNRYKMDLFEFGDEFGMARCPDAAHANVALPADG 180
 QY 181 KLYSATYTDFAIDAVIYRSIGESPTLRVYKDSKMLKEPFYQAVDYGDIYEFREIA 240
 DB 181 KLYSATYTDFAIDAVIYRSIGESPTLRVYKDSKMLKEPFYQAVDYGDIYEFREIA 240
 QY 241 VEYNTMGKVPFPPRAVQVCKNDMGSGORVLEKQWTSFLKARLNCVPGDSHFENILQAVT 300
 DB 241 VEYNTMGKVPFPPRAVQVCKNDMGSGORVLEKQWTSFLKARLNCVPGDSHFENILQAVT 300
 QY 301 DVIRINGRDVVLATFSTPYNSIPGSAVCAVMDLIASVFTGRFKQKSPDSTWTPVDER 360
 DB 301 DVIRINGRDVVLATFSTPYNSIPGSAVCAVMDLIASVFTGRFKQKSPDSTWTPVDER 360
 QY 361 VPKPRPGCCAGSSSLERVATSNRPDDTLNFKTHPLMDEAVPSIENRPNELRTWRYRL 420
 DB 361 VPKPRPGCCAGSSSLERVATSNRPDDTLNFKTHPLMDEAVPSIENRPNELRTWRYRL 420
 QY 421 TKIADVTAAFPYQNTVTVFLGSEKGIILKFLARIGNSGFLNDSFLFEMSVYNSKSCYD 480
 DB 421 TKIADVTAAFPYQNTVTVFLGSEKGIILKFLARIGNSGFLNDSFLFEMSVYNSKSCYD 480
 QY 481 GVEDKRIMGMQJDRASSSLIYVAFSTCVTLNPLGRCEHCKCKCTCIASDPYCGMIKEGG 540
 DB 481 GVEDKRIMGMQJDRASSSLIYVAFSTCVTLNPLGRCEHCKCKCTCIASDPYCGMIKEGG 540
 QY 541 ACSHLSPNRLTFEODIERGNTDGLGDCNSFVALNGHSSSLPFTTSDTAQEGYESR 600
 DB 541 ACSHLSPNRLTFEODIERGNTDGLGDCNSFVALNGHSSSLPFTTSDTAQEGYESR 600
 QY 541 ACSHLSPNRLTFEODIERGNTDGLGDCNSFVALNGHSSSLPFTTSDTAQEGYESR 600
 DB 541 ACSHLSPNRLTFEODIERGNTDGLGDCNSFVALNGHSSSLPFTTSDTAQEGYESR 600
 QY 601 GGMLDKHLIDSPDSTPLGAVSNHNDQKGVIRSYLKQHDQVLPVTLIAVILAVY 660
 DB 601 GGMLDKHLIDSPDSTPLGAVSNHNDQKGVIRSYLKQHDQVLPVTLIAVILAVY 660
 QY 577 -----GVIRSYLKQHDQVLPVTLIAVILAVY 605
 DB 577 -----GVIRSYLKQHDQVLPVTLIAVILAVY 605
 QY 661 MGAVERGITYYCVCDHRKRVAVVQKKEKELTSSRGSMSSVTKLSGLFGDTQSKDPKE 720
 DB 661 MGAVERGITYYCVCDHRKRVAVVQKKEKELTSSRGSMSSVTKLSGLFGDTQSKDPKE 720

DB 606 MGAVERGITYYCVCDHRKRVAVVQKKEKELTSSRGSMSSVTKLSGLFGDTQSKDPKE 665
 QY 721 ALTPLEHNGKLTAPGTAKMLIKADQHDLTALPTPESTPTLQQRKESRGSREWRN 780
 DB 666 ALTPLEHNGKLTAPGTAKMLIKADQHDLTALPTPESTPTLQQRKESRGSREWRN 725
 QY 781 QNLINACTKMPGSPVITPDLPLRASPSHIBSVVLPITQOGYQHEVVOPKMSVYAQ 840
 DB 726 QNLINACTKMPGSPVITPDLPLRASPSHIBSVVLPITQOGYQHEVVOPKMSVYAQ 785
 QY 841 MALEDQAATLEKTKIKELLSKSPNHGYNLVEINDSLPEPKYQREASLGPASLSQTL 900
 DB 786 MALEDQAATLEKTKIKELLSKSPNHGYNLVEINDSLPEPKYQREASLGPASLSQTL 845
 QY 901 SKRIEMHSSSYGVYKRSYPTNSLTRSHQATTLKRNNTSSNSGHLSPNOSFGNDNP 960
 DB 846 SKRIEMHSSSYGVYKRSYPTNSLTRSHQATTLKRNNTSSNSGHLSPNOSFGNDNP 905
 QY 961 PAPQFVDSIQVHSSQPSQAVTVSFQPSLNAVNSLTRSGLKRTPELKPVPKPSFAPS 1020
 DB 906 PAPQFVDSIQVHSSQPSQAVTVSFQPSLNAVNSLTRSGLKRTPELKPVPKPSFAPS 965
 QY 1021 TSMKPNDACT 1030
 DB 966 TSMKPNDACT 975

RESULT 3
 AAM64221
 ID AAM64221 standard; Protein; 974 AA.

AC AAM64221;

DT 06-OCT-1998 (first entry)

DE Human secreted protein from clone CUI45_1.

XX Secreted protein; human fetal brain; nutrition; cytokine; stimulant;
 KW cell proliferation/differentiation; immune system; suppressor; ligand;
 KW regulator; hematopoiesis; tissue growth; activin; inhibin; haemostatic;
 KW chemotaxis; chemokinetic; thrombosis; receptor; cadherin; tumour;
 KW anti-inflammatory.

OS Homo sapiens.

PN M09827205-A2.

XX 25-JUN-1998.

PD 17-DEC-1997; 97WO-US23330.

PF 16-DEC-1997; 97US-0991872.

PR 18-DEC-1996; 96US-0769192.

ER 13-JAN-1997; 97US-0783401.

XX (GENEY) GENETICS INST INC.

PA Agostino MJ, Jacobs K, Lavalie ER, McCoy JM, Weinberg D;
 PI Racie LA, Spaulding V, Treacy M;

DR WPI; 1998-362774/31.

DR N-PSDB; AAV44295.

XX New polynucleotides and secreted proteins - obtained from human
 PT foetal brain, human adult testes, human adult brain and human adult
 PT salivary gland cDNA libraries

PS Claim 17; Page 71-74; 110pp; English.

CC This sequence represents a novel secreted protein from clone CUI45_1
 CC isolated from a human fetal brain cDNA library. This protein has
 CC applications for nutritional use, cytokine and cell
 CC proliferation/differentiation activity, immune stimulating or

suppressing activity, hematopoiesis regulating activity, tissue growth
CC activity, activin/inhibin activity, chemotactic/chemokinetic activity,
CC haemostatic and thrombotic activity, receptor/ligand activity,
CC anti-inflammatory activity, cadherin/tumour invasion suppressor activity,
CC tumour inhibition activity and other activities.

XX Sequence 974 AA;

Query Match 93.4%; Score 5088; DB 19; Length 974;
Best Local Similarity 94.3%; Pred. No. 0;
Matches 971; Conservative 0; Mismatches 3; Indels 56; Gaps 2;

1 MRSBALLYFTLLHFAAGPEDESEPISSHNATKQYVVFVGHKPGNTTQRHLDIOM 60
1 MRSBALLYFTLLHFAAGPEDESEPISSHNATKQYVVFVGHKPGNTTQRHLDIOM 60
61 IMVNGTLYAARDHIYVDIDTSHTEETYSKLTWKSRADVDTCRMKGKDECHNF 120
61 IMVNGTLYAARDHIYVDIDTSHTEETYSKLTWKSRADVDTCRMKGKDECHNF 120
61 IMVNGTLYAARDHIYVDIDTSHTEETYSKLTWKSRADVDTCRMKGKDECHNF 119
121 IKVLLKKNDALFVCGTNAFNPSCNRYKMDLPEFGDEFGQMANCPYDAKANVALPADG 180
120 IKVLLKKNDALFVCGTNAFNPSCNRYKMDLPEFGDEFGQMANCPYDAKANVALPADG 179
181 KLYSATVDFLAIDAVIYRSLGESPTLTVKDKSKLKEPYFVQAVDYGVITYFFFRRIA 240
180 KLYSATVDFLAIDAVIYRSLGESPTLTVKDKSKLKEPYFVQAVDYGVITYFFFRRIA 239
241 VEYNTMGKVPFRRVAVQCNKMDGSOPLKXOMTSFLKARLNCSPGDSHFENLQAVT 300
240 VEYNTMGKVPFRRVAVQCNKMDGSOPLKXOMTSFLKARLNCSPGDSHFENLQAVT 299
301 DVIRINGRDVATSTFPYNSIPGSAVAYMDLADIASVFTGRFKQKSPDSTWTPVDER 360
300 DVIRINGRDVATSTFPYNSIPGSAVAYMDLADIASVFTGRFKQKSPDSTWTPVDER 359
360 VPKPRPGCCAGSSLSERVATSNPPDDTLNFKTHPLMDPAVPSFNNRWFRTWRYRL 419
421 TKIADVTAGPYQNTVTVLGESEKIIILKFLARIGSGFLNLSLFLSEMSVNSEKSYD 480
420 TKIADVTAGPYQNTVTVLGESEKIIILKFLARIGSGFLNLSLFLSEMSVNSEKSYD 479
481 GVEDKRMQDRASSSLYVAFSTCVIYVPLGRCHGCKCTCIASSDPYCGWIKKCG 540
480 GVEDKRMQDRASSSLYVAFSTCVIYVPLGRCHGCKCTCIASSDPYCGWIKKCG 539
541 ACSHLSPNRSLTEQDIERNNTDGLDCHNSFYALN----- 575
540 ACSHLSPNRSLTEQDIERNNTDGLDCHNSFYALN----- 575
601 GGMIDMKHLDPDSDPLGAVSSHNQDKGVIRSESYLKGHDOLPVTLIAVILAFV 660
576 -----GVIRSESYLKGHDOLPVTLIAVILAFV 604
661 MGAVFSGITVYCVCDHRKDVAVVORKEKELTHSRGSMSSVTKLSGLFSDTQSKDPKE 720
605 MGAVFSGITVYCVCDHRKDVAVVORKEKELTHSRGSMSSVTKLSGLFSDTQSKDPKE 664
721 AILTPMHNGKLTAPNTAKMLIADQHHDLTLPTPESTPTLQQRKRSRGRERN 780
665 AILTPMHNGKLTAPNTAKMLIADQHHDLTLPTPESTPTLQQRKRSRGRERN 724
781 QNLINACTKMPMPGSPVITPDLPLRASPSHSIPSVVLLPTIQGYQHEVVDQPMSEVAQ 840
725 QNLINACTKMPMPGSPVITPDLPLRASPSHSIPSVVLLPTIQGYQHEVVDQPMSEVAQ 784
841 MALDQQAATLEFKTIKEHLSKSPNMGVNTVENLDSLPKYVQGEASLGPAGASLSQTGL 900
785 MALDQQAATLEFKTIKEHLSKSPNMGVNTVENLDSLPKYVQGEASLGPAGASLSQTGL 844
901 SKRLEMHSSSYGVYKRSYPTNSLTRSHQATTLKKNNTNSNSHLSRNGSFGRGNDP 960

DB 845 SKRLEMHSSSYGVYKRSYPTNSLTRSHQATTLKKNNTNSNSHLSRNGSFGRGNDP 904
QY 961 PAPQVRDTSIQVWSSSQSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPPVPEKSPAPLS 1020
DB 905 PAPQVRDTSIQVWSSSQSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPPVPEKSPAPLS 964
QY 1021 TSMKENDACT 1030
DB 965 TSMKENDACT 974

RESULT 4

ID AAB23030 standard; Protein: 939 AA.

AC AAB23030;

DT 16-JAN-2001 (first entry)

DB Human semaphorin protein-like splice variant, SECX 2864933-1.

XX SECX protein; human; secreted; membrane-associated; cancer;
XX proliferation regulator; differentiation regulator; non-malignant tumour;
XX immune disorder; autoimmune disease; transplant rejection; allergy; AIDS;
XX infection; inflammatory disorder; arthritis; haematopoietic disorder;
XX skin disorder; cardiovascular disorder; atherosclerosis; restenosis;
XX neurological disease; Alzheimer's disease; trauma; woundings;
XX spinal cord injury; skeletal disorder; cytostatic; immunosuppressive;
XX anti-HIV; anti-inflammatory; antiarthritic; antiarteriosclerotic;
XX neuroprotective; vulnerrary; antiallergic; antimicrobial; cardiant;
XX dermatological; gene therapy.

OS Homo sapiens.

PN WO20053742-A2.

PD 14-SEP-2000.

PF 09-MAR-2000; 2000WO-US06280.

PR 09-MAR-1999; 990US-0123667.

PR 08-MAR-2000; 2000US-0123667.

PA (CURA-) CURAGEN CORP.

PI Shimkets RA;

DR WPI; 2000-594318/56.

DR N-PSDB; AAA93617.

PT Novel human membrane associated or secreted polypeptides and
PT polynucleotides useful for diagnosis, prevention and treatment of
PT pathological states such as cancer, immune, cardiovascular and
PT neurological disorders

PS Claim 1; Fig 2; 151pp; English.

XX Sequences AAB23029-B23048 represent human SECX proteins. The SECX
CC proteins of the invention are either secreted or membrane-associated
CC proteins and act as regulator of cellular proliferation and
CC differentiation. SECX proteins or nucleotides are useful for diagnosing
CC the presence of, or predisposition to, a disease associated with altered
CC levels of SECX proteins and nucleotides. The SECX proteins are also
CC useful to screen compounds that modulate SECX activity or expression. The
CC interaction of a SECX protein with other cellular proteins may be useful
CC to modulate the activity of a partner protein. SECX nucleotides are useful
CC for the recombinant expression and cell survival. SECX nucleotides are useful
CC for the recombinant expression of SECX protein, and may be used detect
CC SECX mRNA or genetic lesions in the SECX gene. They may also be used to
CC modulate SECX expression (e.g., using antisense oligonucleotides). SECX
CC nucleic acid sequences are also useful for identifying a cell or tissue
CC type in a biological sample, and in forensic biology. SECX primers or

CC probes are useful for detecting the presence of SECX nucleotides and for
 CC screening tissue cultures for contamination. Diseases that may be treated
 CC or prevented using SECX proteins or nucleotides include cancer (e.g.,
 CC colorectal carcinoma, prostate cancer), benign tumours, immune disorders
 CC (including autoimmune diseases, transplant rejection, allergies, AIDS),
 CC infections, inflammatory disorders, arthritis, haemopoietic disorders,
 CC skin disorders, cardiovascular disorders, atherosclerosis, restenosis,
 CC neurological diseases (e.g., Alzheimer's disease), trauma (e.g.,
 CC surgical or traumatic wounds, spinal cord injury), and skeletal
 CC disorders.

XX Sequence 939 AA:

Query Match 90.3%; Score 4921; DB 21; Length 939;

Best Local Similarity 99.6%; Pred. No. 0;

Matches 928; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 MRSEALLLYTLHRAAGAFPPDSEPISSHGNTKQYVFGHKGGRNTTORHLDIQM 60
DB 1 MRSEALLLYTLHRAAGAFPPDSEPISSHGNTKQYVFGHKGGRNTTORHLDIQM 60
QY 61 IMNMGTLIARADHIYTVDTIDTSHTEETIYCSKLTWKSROADVDTCKMKGKHDECHNF 120
DB 61 IMNMGTLIARADHIYTVDTIDTSHTEETIYCSKLTWKSROADVDTCKMKGKHDECHNF 120
QY 121 IKVLLKQNDALFVCGTNAFNPSCRNYKMDLEPFDEFSGNARCPYDAKXANVALFADG 180
DB 121 IKVLLKQNDALFVCGTNAFNPSCRNYKMDLEPFDEFSGNARCPYDAKXANVALFADG 180
QY 181 KLYSAVTVPLAIDANTYISLGSPTLRVYKXSKLKEPYQVADYDGYIFFPRETA 240
DB 181 KLYSAVTVPLAIDANTYISLGSPTLRVYKXSKLKEPYQVADYDGYIFFPRETA 240
QY 241 VERNMGKVPFPAVAVOCNMDMGSOQVLEKQWTSFLKRLNCSVPGDSHFYENILQAVT 300
DB 241 VERNMGKVPFPAVAVOCNMDMGSOQVLEKQWTSFLKRLNCSVPGDSHFYENILQAVT 300
QY 301 DVIRINGRDVLTATFSTPYNSIPGSAVCAVMDLIVSFTGRFKEKSPDSTWTPVDER 360
DB 301 DVIRINGRDVLTATFSTPYNSIPGSAVCAVMDLIVSFTGRFKEKSPDSTWTPVDER 360
QY 361 VPPRRGCCGSSSLERATSNFPDDTNFITHLMDEAVSIFNRPFELTMYRIL 420
DB 361 VPPRRGCCGSSSLERATSNFPDDTNFITHLMDEAVSIFNRPFELTMYRIL 420
QY 421 TKTAVIDTAAGPYQNHVTFLGSEKGIILKFLARIGNSGFLNDSLPLEMSVNSEKSYD 480
DB 421 TKTAVIDTAAGPYQNHVTFLGSEKGIILKFLARIGNSGFLNDSLPLEMSVNSEKSYD 480
QY 481 GVEDKRIWQOLDRASSSLIVAFSTCVIKVPLGRCERHCKCKTCTASRPYCGMITEGG 540
DB 481 GVEDKRIWQOLDRASSSLIVAFSTCVIKVPLGRCERHCKCKTCTASRPYCGMITEGG 540
QY 541 ACSHLASNSRLTEODIERGNTDGLGCHNSFVALNHSLSLPTSTTSSTAOEGYESR 600
DB 541 ACSHLASNSRLTEODIERGNTDGLGCHNSFVALNHSLSLPTSTTSSTAOEGYESR 600
QY 601 GGMMDWHLDSPESTPLGAVSSHNQDKGVIRESYLKGHDQVETLLAIAVILAFV 660
DB 601 GGMMDWHLDSPESTPLGAVSSHNQDKGVIRESYLKGHDQVETLLAIAVILAFV 660
QY 661 MGAVFSGITVYCVCDHRKDVAVVQRKEKELTHSRSGSSSVTLKSLGFGDTQKDPPE 720
DB 661 MGAVFSGITVYCVCDHRKDVAVVQRKEKELTHSRSGSSSVTLKSLGFGDTQKDPPE 720
QY 721 AILPLIMNGLATPAGTAKMLIKADQHLHLTLPTPESTPTLOQKRRKPSRGSEMERN 780
DB 721 AILPLIMNGLATPAGTAKMLIKADQHLHLTLPTPESTPTLOQKRRKPSRGSEMERN 780
QY 781 ONLINACTCKMPKMGSPVITPDLPLRASPSHIPSVVVLPTQGYOHEVYDQPMSEVAQ 840
DB 781 ONLINACTCKMPKMGSPVITPDLPLRASPSHIPSVVVLPTQGYOHEVYDQPMSEVAQ 840

```

```

QY 841 MALEDQATLEKYTKIKELSSKSPNHCNVLENLDSUPPKVPOREASIGPPGASLSQTGL 900
DB 841 MALEDQATLEKYTKIKELSSKSPNHCNVLENLDSUPPKVPOREASIGPPGASLSQTGL 900
QY 901 SKRLEMHSSSYGVYKRSYPTNSLTNSHOAT 932
DB 901 SKRLEMHSSSYGVYKRSYPTNSLTNSHOAT 932

```

RESULT 5

AAAB23031 ID AAAB23031 standard; Protein, 884 AA.

AAAB23031; 16-JAN-2001 (first entry)

Human semaphorin protein-like splice variant, SECX 2864933-2.

SECX protein; human; secreted; membrane-associated; cancer;
 proliferation regulator; differentiation regulator; non-malignant tumour;
 immune disorder; autoimmune disease, transplant rejection, allergy; AIDS;
 infection; inflammatory disorder; arthritis, haemopoietic disorder;
 skin disorder; cardiovascular disorder; atherosclerosis, restenosis;
 neurological disease; Alzheimer's disease; trauma; wounding;
 spinal cord injury; skeletal disorder; cytostatic; immunosuppressive;
 anti-HIV; antiinflammatory; antiarthritic; antiatherosclerotic;
 neuroprotective; vulnerability; antiallergic; antimicrobial; cardiant;
 dermatological; gene therapy.

Homo sapiens.

WO200057742-A2.

14-SEP-2000.

09-MAR-2000; 2000WO-US06280.

09-MAR-1999; 99US-0123667.
 08-MAR-2000; 2000US-0123667.

(CURA-) CURAGEN CORP.

Shinketsu RA;

WPI, 2000-594318/56.

N-PSDB; AAA93618.

Novel human membrane associated or secreted polypeptides and
 polynucleotides useful for diagnosis, prevention and treatment of
 pathological states such as cancer, immune, cardiovascular and
 neurological disorders -

Claim 1; Fig 3; 15pp; English.

Sequences AAAB23029-B23048 represent human SECX proteins. The SECX
 proteins and act as regulator of cellular proliferation and
 differentiation. SECX proteins or nucleotides are useful for diagnosing
 the presence of, or predisposition to, a disease associated with altered
 levels of SECX proteins and nucleotides. The SECX proteins are also
 useful to screen compounds that modulate SECX activity or expression. The
 interaction of a SECX protein with other cellular proteins may be useful
 to modulate the activity of a partner protein. SECX nucleotides are useful
 for the recombinant expression of SECX protein, and may be used to detect
 SECX mRNA or genetic lesions in the SECX gene. They may also be used to
 modulate SECX expression (e.g., using antisense oligonucleotides). SECX
 nucleic acid sequences are also useful for identifying a cell or tissue
 type in a biological sample, and in forensic biology. SECX primers or
 probes are useful for detecting the presence of SECX nucleotides and for
 screening tissue cultures for contamination. Diseases that may be treated
 or prevented using SECX proteins or nucleotides include cancer (e.g.,

CC colorectal carcinoma, prostate cancer), benign tumours, immune disorders
 CC (including autoimmune diseases, transplant rejection, allergies, AIDS),
 CC infections, inflammatory disorders, arthritis, haematopoietic disorders,
 CC skin disorders, cardiovascular disorders, atherosclerosis, restenosis,
 CC neurological diseases (e.g., Alzheimer's disease), trauma (e.g.,
 CC surgical or traumatic wounds, spinal cord injury), and skeletal
 CC disorders.

XX Sequence 884 AA;

Query Match 84.2%; Score 4590.5; DB 21; Length 884;

Best Local Similarity 93.7%; Pred. No. 0; Mismatches 2; Indels 55; Gaps 1;

Matches 873; Conservative 2; Mismatches 2; Indels 55; Gaps 1;

QY 1 MRSBALLYFLTFAGAGPEDESEPISSHGNTYKQYVFGHKPGRNTTQRRLDIOM 60
 DB 1 MRSBALLYFLTFAGAGPEDESEPISSHGNTYKQYVFGHKPGRNTTQRRLDIOM 60
 QY 61 IMINMGTLTYAARDHIYTVDDIDTSHTEBIYCSKLTWKSROADVDTGMKXKDECHNF 120
 DB 61 IMINMGTLTYAARDHIYTVDDIDTSHTEBIYCSKLTWKSROADVDTGMKXKDECHNF 120
 QY 121 IKULKKRNDLALFYCGNNAENPSCNRYMDLLEPGDFSGARCPYDAKIANVALFADG 180
 DB 121 IKULKKRNDLALFYCGNNAENPSCNRYMDLLEPGDFSGARCPYDAKIANVALFADG 180
 QY 121 IKVLIKKNDLALFYCGNNAENPSCNRYMDLLEPGDFSGARCPYDAKIANVALFADG 180
 DB 121 IKVLIKKNDLALFYCGNNAENPSCNRYMDLLEPGDFSGARCPYDAKIANVALFADG 180
 QY 181 KLYSATVDFLAIDAVIYRSLGSEPTLTVMKDSKMLKEPYFVQAVDYGYIYFFPREIA 240
 DB 181 KLYSATVDFLAIDAVIYRSLGSEPTLTVMKDSKMLKEPYFVQAVDYGYIYFFPREIA 240
 QY 181 KLYSATVDFLAIDAVIYRSLGSEPTLTVMKDSKMLKEPYFVQAVDYGYIYFFPREIA 240
 DB 181 KLYSATVDFLAIDAVIYRSLGSEPTLTVMKDSKMLKEPYFVQAVDYGYIYFFPREIA 240
 QY 241 VEYNTGKVYPPRAQVCKNDWGSGQVLEKQWTSFLKARLNCSPGSHHYFNLLQAVT 300
 DB 241 VEYNTGKVYPPRAQVCKNDWGSGQVLEKQWTSFLKARLNCSPGSHHYFNLLQAVT 300
 QY 241 VEYNTGKVYPPRAQVCKNDWGSGQVLEKQWTSFLKARLNCSPGSHHYFNLLQAVT 300
 DB 241 VEYNTGKVYPPRAQVCKNDWGSGQVLEKQWTSFLKARLNCSPGSHHYFNLLQAVT 300
 QY 301 DVIRINGRDVLTATSTPNYSIPGSAVCAVMDLJASVFTGFRKQKSPDSTWTTPVDER 360
 DB 301 DVIRINGRDVLTATSTPNYSIPGSAVCAVMDLJASVFTGFRKQKSPDSTWTTPVDER 360
 QY 301 DVIRINGRDVLTATSTPNYSIPGSAVCAVMDLJASVFTGFRKQKSPDSTWTTPVDER 360
 DB 301 DVIRINGRDVLTATSTPNYSIPGSAVCAVMDLJASVFTGFRKQKSPDSTWTTPVDER 360
 QY 361 VKPRPGCCAGSSSLERVATSNPEPDDTLNFKTHPLMDEAVPSIFNRPWELRWRYRL 420
 DB 361 VKPRPGCCAGSSSLERVATSNPEPDDTLNFKTHPLMDEAVPSIFNRPWELRWRYRL 420
 QY 361 VKPRPGCCAGSSSLERVATSNPEPDDTLNFKTHPLMDEAVPSIFNRPWELRWRYRL 420
 DB 361 VKPRPGCCAGSSSLERVATSNPEPDDTLNFKTHPLMDEAVPSIFNRPWELRWRYRL 420
 QY 421 TKIAVDYAAQPYQHTVVFVLSGEGKILKFLARIGNSGLNLSLLEMSVYNSKSCYD 480
 DB 421 TKIAVDYAAQPYQHTVVFVLSGEGKILKFLARIGNSGLNLSLLEMSVYNSKSCYD 480
 QY 421 TKIAVDYAAQPYQHTVVFVLSGEGKILKFLARIGNSGLNLSLLEMSVYNSKSCYD 480
 DB 421 TKIAVDYAAQPYQHTVVFVLSGEGKILKFLARIGNSGLNLSLLEMSVYNSKSCYD 480
 QY 481 GVEDRIRIMQMDRASSSLIYVAFSTCVIPLGRGCRHKCKKTCIASRDPYCGMIKESG 540
 DB 481 GVEDRIRIMQMDRASSSLIYVAFSTCVIPLGRGCRHKCKKTCIASRDPYCGMIKESG 540
 QY 481 GVEDRIRIMQMDRASSSLIYVAFSTCVIPLGRGCRHKCKKTCIASRDPYCGMIKESG 540
 DB 481 GVEDRIRIMQMDRASSSLIYVAFSTCVIPLGRGCRHKCKKTCIASRDPYCGMIKESG 540
 QY 541 ACSHTSPNSRLTFEODIERGNTDGLDCHNSFVALNGHSSLLPSTTTSDSTAQGYESR 600
 DB 541 ACSHTSPNSRLTFEODIERGNTDGLDCHNSFVALNGHSSLLPSTTTSDSTAQGYESR 600
 QY 541 ACSHTSPNSRLTFEODIERGNTDGLDCHNSFVALNGHSSLLPSTTTSDSTAQGYESR 600
 DB 541 ACSHTSPNSRLTFEODIERGNTDGLDCHNSFVALNGHSSLLPSTTTSDSTAQGYESR 600
 QY 601 GGNLDMKHLIDBPSDIDPLGAVSHHOKKQKVIKRSYKKGHDQVAVPTLLAIVLAVFV 660
 DB 601 GGNLDMKHLIDBPSDIDPLGAVSHHOKKQKVIKRSYKKGHDQVAVPTLLAIVLAVFV 660
 QY 601 GGNLDMKHLIDBPSDIDPLGAVSHHOKKQKVIKRSYKKGHDQVAVPTLLAIVLAVFV 660
 DB 601 GGNLDMKHLIDBPSDIDPLGAVSHHOKKQKVIKRSYKKGHDQVAVPTLLAIVLAVFV 660
 QY 661 MGAVSGITVYCVCDHRRKDVAVVORKEKELTSSRGSSSVTKLSGLFGDTQSKDPKE 720
 DB 661 MGAVSGITVYCVCDHRRKDVAVVORKEKELTSSRGSSSVTKLSGLFGDTQSKDPKE 720
 QY 661 MGAVSGITVYCVCDHRRKDVAVVORKEKELTSSRGSSSVTKLSGLFGDTQSKDPKE 720
 DB 661 MGAVSGITVYCVCDHRRKDVAVVORKEKELTSSRGSSSVTKLSGLFGDTQSKDPKE 720
 QY 721 AITPLPMNGKATPNTAKMLIKADQHLIDTALPTPESTPTLQOKRPSRGRSWERN 780
 DB 721 AITPLPMNGKATPNTAKMLIKADQHLIDTALPTPESTPTLQOKRPSRGRSWERN 780
 QY 721 AITPLPMNGKATPNTAKMLIKADQHLIDTALPTPESTPTLQOKRPSRGRSWERN 780
 DB 721 AITPLPMNGKATPNTAKMLIKADQHLIDTALPTPESTPTLQOKRPSRGRSWERN 780
 QY 781 QNLINACTKMPMGSPVPTDPLASPSHISVVVLTPTQGYGHEVYDOPKMEVQA 840
 DB 781 QNLINACTKMPMGSPVPTDPLASPSHISVVVLTPTQGYGHEVYDOPKMEVQA 840
 QY 781 QNLINACTKMPMGSPVPTDPLASPSHISVVVLTPTQGYGHEVYDOPKMEVQA 840
 DB 781 QNLINACTKMPMGSPVPTDPLASPSHISVVVLTPTQGYGHEVYDOPKMEVQA 840
 QY 841 MALEDOAAITLEYKIKHLSKSPNGVNLVENLDSLPKVPVPREASLGPAGASLSQGT 900
 DB 841 MALEDOAAITLEYKIKHLSKSPNGVNLVENLDSLPKVPVPREASLGPAGASLSQGT 900
 QY 786 MALEDOAAITLEYKIKHLSKSPNGVNLVENLDSLPKVPVPREASLGPAGASLSQGT 900
 DB 786 MALEDOAAITLEYKIKHLSKSPNGVNLVENLDSLPKVPVPREASLGPAGASLSQGT 900

QY 901 SKRLEMHSSVGVYDKRSPYPTNSLTRSHOAT 932
 DB 846 SKRLEMHSSVGVYDKRSPYPTNSLTRSHOAT 877

RESULT 6
 AAB95139
 ID AAB95139 standard; Protein; 699 AA.

AC AAB95139;
 XX 26-JUN-2001 (first entry)
 DT 26-JUN-2001 (first entry)
 DE Human protein sequence SEQ ID NO:17154.

DE Human protein sequence SEQ ID NO:17154.
 DE Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX Homo sapiens.

XX Homo sapiens.
 XX EPI074617-A2.
 XX 07-FEB-2001.

XX 28-JUL-2000; 2000CEP-0116126.
 XX 29-JUL-1999; 99JP-0248036.
 XX 27-AUG-1999; 99JP-0300253.
 XX 11-JAN-2000; 2000JP-018776.
 XX 02-MAY-2000; 2000JP-0183767.
 XX 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.
 XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -

XX Claim 8; SEQ ID 17154; 2537bp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification, where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides; and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.

XX Sequence 699 AA;
 SQ Query Match 67.5%; Score 3677; DB 22; Length 699;
 Best Local Similarity 100.0%; Pred. No. 1.6e-310;

Matches	699;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	332	MLDIASVFTGPEKEKSPDSTWTPVDESEVPPKPPCCAGSSSLERATSNFPPDTLNF	391						
Db	1	MLDIASVFTGPEKEKSPDSTWTPVDESEVPPKPPCCAGSSSLERATSNFPPDTLNF	60						
QY	392	ITHTPLMDEAVSIFNRPWFELRTWRYRLTKIAVDTAAGPYQNHVYFEGSKGIILKFL	451						
Db	61	IXTHPLMDEAVSIFNRPWFELRTWRYRLTKIAVDTAAGPYQNHVYFEGSKGIILKFL	120						
QY	452	AATGNSGFINDSLFLEEMSVYNSEKCSYDVEDKRLMGQLDRASSSLVYASTCIVKP	511						
Db	121	ARIQNSGFINDSLFLEEMSVYNSEKCSYDVEDKRLMGQLDRASSSLVYASTCIVKP	180						
QY	512	LGRCEHSGCKTKCIASRDPCYCGWIKGAGCSHLSNSRLTEODIERGNTGLGDCHNS	571						
Db	181	LGRCEHSGCKTKCIASRDPCYCGWIKGAGCSHLSNSRLTEODIERGNTGLGDCHNS	240						
QY	572	FVALNGHSSSLPSTTTSSTAOEGESRGMLDMGHLSDPSITDPLGAVSSHNDKK	631						
Db	241	FVALNGHSSSLPSTTTSSTAOEGESRGMLDMGHLSDPSITDPLGAVSSHNDKK	300						
QY	632	GVTRSEYLVGHQDQVAVTLLAIVLLAFVWGAVERSGITVYCVCDHRKDVAVVQREKEL	691						
Db	301	GVTRSEYLVGHQDQVAVTLLAIVLLAFVWGAVERSGITVYCVCDHRKDVAVVQREKEL	360						
QY	692	THSRGSSSVTLKSLGFGDTQSKDPKPEAILPLMHNGKLATPNTAKYLIRADQHHLD	751						
Db	361	THSRGSSSVTLKSLGFGDTQSKDPKPEAILPLMHNGKLATPNTAKYLIRADQHHLD	420						
QY	752	LVALPPESTPTVQQRKPSRGSEMERONLINACTCKMPKPSPIPTDPLRASPSH	811						
Db	421	LVALPPESTPTVQQRKPSRGSEMERONLINACTCKMPKPSPIPTDPLRASPSH	480						
QY	812	IPSVVVLPIITQGGYQHEHYVQPKMSEVAQALDEQATLEKYITKEHLSKSPNHYV	871						
Db	481	IPSVVVLPIITQGGYQHEHYVQPKMSEVAQALDEQATLEKYITKEHLSKSPNHYV	540						
QY	872	ENLDSLPPKYPORASLGGPPGASLSQTGLSKRLMHSSSYGVYKRSYTNLSITRSHQ	931						
Db	541	ENLDSLPPKYPORASLGGPPGASLSQTGLSKRLMHSSSYGVYKRSYTNLSITRSHQ	600						
QY	932	TYLKRNTSSNSGSHSRNOSFGGNDPPPPAPQVDSIQVHSSQPSQAVTVSRQPSLNA	991						
Db	601	TYLKRNTSSNSGSHSRNOSFGGNDPPPPAPQVDSIQVHSSQPSQAVTVSRQPSLNA	660						
QY	992	YNSLTRSGIKRTPSLKDPVPEKPSFAPLSTSMKENDACT	1030						
Db	661	YNSLTRSGIKRTPSLKDPVPEKPSFAPLSTSMKENDACT	699						

RESULT 7
AAB23043
ID AAB23043 standard; Protein; 630 AA.

XX AAB23043;
XX
XX 16-JAN-2001 (first entry)

DE Human semaphorin protein-like splice variant, SECK pCR2.1-2864933.

XX SECK protein; human; secreted; membrane-associated; cancer;
XX Proliferation regulator; differentiation regulator; non-malignant tumour;
XX immune disorder; autoimmune disease; transplant rejection; allergy; AIDS;
XX infection; inflammatory disorder; arthritis; haematopoietic disorder;
XX skin disorder; cardiovascular disorder; atherosclerosis; restenosis;
XX neurological disease; Alzheimer's disease; trauma; wounding;
XX spinal cord injury; skeletal disorder; cytostatic; immunosuppressive;
XX anti-HIV; anti-inflammatory; antiatheritic; antiarteriosclerotic;
XX neuroprotective; vulnery; antiallergic; antimicrobial; cardiac;
XX dermatological; gene therapy.

OS Homo sapiens.

XX	WO200053742-A2.
XX	14-SEP-2000.
XX	09-MAR-2000; 2000WO-US06280.
XX	09-MAR-1999; 99US-0123667.
XX	08-MAR-2000; 2000US-0123667.
XX	(CURAGEN CORP.
XX	Shinketsu RA;
XX	WPI: 2000-594318/56.
XX	N-PSDB; AAA93630.
XX	Novel human membrane associated or secreted polypeptides and
XX	polynucleotides useful for diagnosis, prevention and treatment of
XX	pathological states such as cancer, immune, cardiovascular and
XX	neurological disorders
XX	Claim 1; Fig 15; 15pp; English.
XX	Sequences AAB23029-B23048 represent human SECK proteins. The SECK
XX	proteins of the invention are either secreted or membrane-associated
XX	and act as regulator of cellular proliferation and
XX	differentiation. SECK proteins or nucleotides are useful for diagnosing
XX	the presence of, or predisposition to, a disease associated with altered
XX	levels of SECK proteins and nucleotides. The SECK proteins are also
XX	useful to screen compounds that modulate SECK activity or expression. The
XX	interaction of a SECK protein with other cellular proteins may be useful
XX	to modulate the activity of a partner protein. SECK nucleotides are useful
XX	cellular differentiation and cell survival. SECK nucleotides are useful
XX	for the recombinant expression of SECK gene. They may also be used to
XX	SECK mRNA or genetic lesions in the SECK gene. They may also be used to
XX	modulate SECK expression (e.g., using antisense oligonucleotides). SECK
XX	nucleic acid sequences are also useful for identifying a cell or tissue
XX	type in a biological sample, and in forensic biology. SECK primers or
XX	probes are useful for detecting the presence of SECK nucleotides and for
XX	screening tissue cultures for contamination. Diseases that may be treated
XX	or prevented using SECK proteins or nucleotides include cancer (e.g.,
XX	colorectal carcinoma, prostate cancer), benign tumours, immune disorders
XX	(including autoimmune diseases, transplant rejection, allergies, AIDS),
XX	infections, inflammatory disorders, arthritis, haematopoietic disorders,
XX	skin disorders, cardiovascular disorders, atherosclerosis, restenosis,
XX	neurological diseases (e.g., Alzheimer's disease), trauma (e.g.,
XX	surgical or traumatic wounds, spinal cord injury), and skeletal
XX	disorders.
XX	Sequence 630 AA;
XX	Query Match 61.9%; Score 3373; DB 21; Length 630;
XX	Best Local Similarity 99.8%; Pred. No. 3.8e-284;
XX	Matches 628; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY	17 GAGPPEDESEPISSHGNTKYQYPVVGKPKGRNTTQRRHLDIQMIMTNGTLYIARDDHI 76
Db	1 GSGPPEDESEPISSHGNTKYQYPVVGKPKGRNTTQRRHLDIQMIMTNGTLYIARDDHI 60
QY	77 YTVDDITSHTEIIVCSKLLTWKSRQADVDYTRMGKHKDECHNFIYLLKKNDALFVCG 136
Db	61 YTVDDITSHTEIIVCSKLLTWKSRQADVDYTRMGKHKDECHNFIYLLKKNDALFVCG 120
QY	137 TNAFNPSCRNYKMTLPEFGDEFGSMARCPYDAKHANVALFADGKLSATVTDPLAIDAV 196
Db	121 TNAFNPSCRNYKMTLPEFGDEFGSMARCPYDAKHANVALFADGKLSATVTDPLAIDAV 180
QY	197 IYRSIGSEPTLRATVYHDSKMLKEPFYQAVDYGDIYFFPELAVENYTMGVYPRVAQ 256
Db	181 IYRSIGSEPTLRATVYHDSKMLKEPFYQAVDYGDIYFFPELAVENYTMGVYPRVAQ 240
QY	257 VCKNDMGSGQRVLEKQWTFPLKARLNGVPSGDSHRYFNLQAVTVIRNGRDVLAIFS 316


```

Db      241 VCNNDGSGQRLVEKQMTSLKARLNGSVGDHSHFYNNILQAVTDVLRINGRDVIATRS 300
Qy      317 TPVNSIFGSAVCAYMDLDSVFTGRKEQKSPDSTMTVPDPRRVPKRRPGCCAGSSSLR 376
Db      301 TPVNSIFGSAVCAYMDLDSVFTGRKEQKSPDSTMTVPDPRRVPKRRPGCCAGSSSLR 360
Qy      377 RYANSNFPDPTLNFITHTPLMDRAVPSIFNRPMFLRTMYRLLTKIADVTAAAGPYONHT 436
Db      361 RYANSNFPDPTLNFITHTPLMDRAVPSIFNRPMFLRTMYRLLTKIADVTAAAGPYONHT 420
Qy      437 VFLGSEKGIILKFLARIGNSGFNLDSLPLEMSVYNSEKSYDGVEDKRIKMGQIDRAS 496
Db      421 VFLGSEKGIILKFLARIGNSGFNLDSLPLEMSVYNSEKSYDGVEDKRIKMGQIDRAS 480
Qy      497 SSLVAFSTCVIKVPLGRCERHCKCKTICASRPYCGMKKEGACSHLSPNSRLTEOD 556
Db      481 SSLVAFSTCVIKVPLGRCERHCKCKTICASRPYCGMKKEGACSHLSPNSRLTEOD 540
Qy      557 IERNTDGLGCHNSFVALNGHSSSLPSTTSDTAQEGYERSGMLDMKHLDSPDST 616
Db      541 IERNTDGLGCHNSFVALNGHSSSLPSTTSDTAQEGYERSGMLDMKHLDSPDST 600
Qy      617 DPLGAVSSHNHODKKGVIRESYLKGHDQL 645
Db      601 DPLGAVSSHNHODKKGVIRESYLKGHDQL 629

```

RESULT 8

AAB95317 standard; Protein; 574 AA.

AAB95317;

26-JUN-2001 (first entry)

Human protein sequence SEQ ID NO:17568.

Human; primer; detection; diagnosis; antisense therapy; gene therapy.

Homo sapiens.

EP1074617-A2.

07-FEB-2001.

28-JUL-2000; 2000EP-0116126.

29-JUL-1999; 99QP-0248036.

27-AUG-1999; 99QP-0300253.

11-JUN-2000; 2000JP-0118776.

02-MAY-2000; 2000JP-0183767.

09-JUN-2000; 2000JP-0241899.

(HELI-) HELIX RES INST.

Oca T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

WPI; 2001-318749/34.

Primer sets for synthesizing polynucleotides, particularly the 5602

full-length CDNA defined in the specification, and for the detection

and/or diagnosis of the abnormality of the proteins encoded by the

full-length CDNA -

Claim 8; SEQ ID 17568; 2537pp + CD ROM; English.

The present invention describes primer sets for synthesizing 5602

oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the combination of oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length CDNA. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length CDNA. The primers allow obtaining of the full-length CDNA easily without any special methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human CDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

Sequence 574 AA;

Query Match 54.7%; Score 2980.5; DB 22; Length 574;

Best Local Similarity 90.7%; Pred. No. 4.6e-250; Indels 59; Gaps 1;

Matches 574; Conservative 0; Mismatches 0;

398 MDEAVPSIFNRPMFLRTMYRLLTKIADVTAAAGPYONHTVFLGSEKGIILKFLARIGNS 457

1 MDEAVPSIFNRPMFLRTMYR-----CSYDGVEDKRIKMGQIDRASSLVAFSTCVIKVPLGRCER 517

458 GFNLDSLPLEMSVYNSEKSYDGVEDKRIKMGQIDRASSLVAFSTCVIKVPLGRCER 517

21 -----CSYDGVEDKRIKMGQIDRASSLVAFSTCVIKVPLGRCER 61

518 HGCKKTCIASRDPYCGMKKEGACSHLSPNSRLTFEODIERGNTDGLGCHNSFVALNG 577

62 HGCKKTCIASRDPYCGMKKEGACSHLSPNSRLTFEODIERGNTDGLGCHNSFVALNG 121

578 HSSSLPSTTSDTAQEGYERSGMLDMKHLDSPDSTPLGAVSSHNHODKKGVIRE 637

122 HSSSLPSTTSDTAQEGYERSGMLDMKHLDSPDSTPLGAVSSHNHODKKGVIRE 181

638 YLKGHDQIVPVTLLAIAVILAFVNGAVNSGITYVCVCHRRKDVAVVORKEKELTHSRG 697

182 YLKGHDQIVPVTLLAIAVILAFVNGAVNSGITYVCVCHRRKDVAVVORKEKELTHSRG 241

698 SMSVYTKLSGLFGDTQSKDPPEALITPLMNGKLTATGNTAKMLIKADQHTLDTALPT 757

242 SMSVYTKLSGLFGDTQSKDPPEALITPLMNGKLTATGNTAKMLIKADQHTLDTALPT 301

758 PESTPTLQOKRKPSRSGREMERONLINACTKMPMGSPVPTDPLPASSHIPSVVV 817

302 PESTPTLQOKRKPSRSGREMERONLINACTKMPMGSPVPTDPLPASSHIPSVVV 361

818 LPITQGGYQHEVYDQPKSEVAQMALEDOAATLEKTKKEHSSKSPHNGVNLVENDSL 877

362 LPITQGGYQHEVYDQPKSEVAQMALEDOAATLEKTKKEHSSKSPHNGVNLVENDSL 421

878 PKVPPREASLGPFGASLSQGTGSKRLKEMHSSSYGVYKRSYPTNLSLRHQATTLKEN 937

422 PKVPPREASLGPFGASLSQGTGSKRLKEMHSSSYGVYKRSYPTNLSLRHQATTLKEN 481

938 NTSSNSHLSRNQSGFGDNPPAPQVDSIQVHSSQPSGQAVVYTSRQSLNAYNSLFR 997

482 NTSSNSHLSRNQSGFGDNPPAPQVDSIQVHSSQPSGQAVVYTSRQSLNAYNSLFR 541

998 SGLKRTPSLKPDVPPKPSFAPLSTSKPPDACT 1030

542 SGLKRTPSLKPDVPPKPSFAPLSTSKPPDACT 574

RESULT 9

AAB94104 standard; Protein; 562 AA.

AAB94104

XX

AC AAB94104;
 XX
 DT 26-JUN-2001 (first entry)
 DE Human protein sequence SEQ ID NO:14328.
 XX
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX Homo sapiens.
 XX EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-0116126.
 XX
 PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 XX (HELI-) HELIX RES INST.
 PA Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 DR WPI; 2001-318749/34.
 XX
 XX Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX
 XX Claim 8; SEQ ID 14328; 2537JP + CD ROW; English.
 XX
 XX The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 CC
 CC Sequence 562 AA;
 SQ
 Query Match 54.2%; Score 2952; DB 22; Length 562;
 Best local similarity 100.0%; Pred. No. 1.4e-247;
 Matches 562; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 589 SDSTAQGEYESSRGMLDWKHLSDPDSTPLGAVSSHNDKKGVIRESLYKGDQLVV 648
 DB 121 SDSTAQGEYESSRGMLDWKHLSDPDSTPLGAVSSHNDKKGVIRESLYKGDQLVV 180
 QY 649 TLDAVIAVIAFVNGAVPSGITYVCVCDHRKRDVAIVQKKEKELTHSRGSMSSVTKLSGL 708
 DB 181 TLDAVIAVIAFVNGAVPSGITYVCVCDHRKRDVAIVQKKEKELTHSRGSMSSVTKLSGL 240
 QY 709 FGDTSQCDPRPEALITPLMNGKLAATPGNTAKMLIKADQHLDLTALPTESPFTLQOKR 768
 DB 241 FGDTSQCDPRPEALITPLMNGKLAATPGNTAKMLIKADQHLDLTALPTESPFTLQOKR 300
 QY 769 KPSRGSREMERNOULINACTKMPMGSPVPIPTDLPASPSHIPSVVILPTQOQYOE 828
 DB 301 KPSRGSREMERNOULINACTKMPMGSPVPIPTDLPASPSHIPSVVILPTQOQYOE 360
 QY 829 YVDPKXSEVAQVMALEDQAATLEYKTIKEHLSSSPNHNQVNLVNLSPKVPQREASL 888
 DB 361 YVDPKXSEVAQVMALEDQAATLEYKTIKEHLSSSPNHNQVNLVNLSPKVPQREASL 420
 QY 889 GPPGASISQTLGSKRLMHHSSSYGVDPYKRSYPTNSLTRSHQATTLKNNNTSSNSHLS 948
 DB 421 GPPGASISQTLGSKRLMHHSSSYGVDPYKRSYPTNSLTRSHQATTLKNNNTSSNSHLS 480
 QY 949 RNOSFGRGNDPPAPORVDSIQVHSSQPSGQAVTVSRPSINAVNSLTRSGIKRTPSLKP 1008
 DB 481 RNOSFGRGNDPPAPORVDSIQVHSSQPSGQAVTVSRPSINAVNSLTRSGIKRTPSLKP 540
 QY 1009 DVPPKPSFAPLSTSMKPNDACT 1030
 DB 541 DVPPKPSFAPLSTSMKPNDACT 562
 RESULT 10
 ID AAM93444
 XX AAM93444 standard; Protein; 562 AA.
 AC AAM93444;
 XX
 DT 06-NOV-2001 (first entry)
 DE Human polypeptide; SEQ ID NO: 3088.
 XX
 KW Human; full length cDNA; cDNA synthesis; oligo-capping.
 XX Homo sapiens.
 XX EP1130094-A2.
 PN 05-SEP-2001.
 PD 07-JUL-2000; 2000EP-0114089.
 PF 08-JUL-1999; 99JP-0194486.
 PR 11-JAN-2000; 2000JP-0118774.
 PR 02-MAY-2000; 2000JP-0183765.
 XX
 XX (HELI-) HELIX RES INST.
 PA Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 DR WPI; 2001-524255/58.
 DR N-PSDB; AAK94365.
 XX
 XX Claim 8; SEQ ID NO 3088; 1380bp + sequence listing; English.
 XX The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been
 CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA

molecules have been determined. Primers for synthesizing the full length cDNA are useful for clarifying the function of the protein encoded by the cDNA. The full length clones were obtained by construction of full length enriched cDNA libraries that were synthesised by the oligo-capping method. The primers enable the production of the full length cDNA easily without any special methods. The present sequence is a polypeptide encoded by a full length human cDNA of the invention.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in CD-ROM format directly from EPO.

Sequence 562 AA:

Query Match 54.1%; Score 2947; DB 22; Length 562;
Best Local Similarity 99.8%; Pred. No. 3.8e-247;
Matches 561; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

CC 469 MSVNSKCSYDGVEDKRIKMGQDRASSSLVAFSTVIVIPGRGGRGCKKCTTAS 528
CC 1 MSVNSKCSYDGVEDKRIKMGQDRASSSLVAFSTVIVIPGRGGRGCKKCTTAS 60
CC 529 RDPYCGWIKGAGCSHLSPLNSRLTFEODIERGNTDGLDCHNSFVALNGHSSLLPSTTT 588
CC 61 RDPYCGWIKGAGCSHLSPLNSRLTFEODIERGNTDGLDCHNSFVALNGHSSLLPSTTT 120
CC 589 SDSTAQEGYSGRGMMDKHLIDPSDTPDPIGAVSSHHNODKKGYTRRSYTKGHDQVLYPV 648
CC 121 SDSTAQEGYSGRGMMDKHLIDPSDTPDPIGAVSSHHNODKKGYTRRSYTKGHDQVLYPV 180
CC 649 TLLAVALIAPFMGAVFSGITVYCVCDHRKDVAVQKKKELTSSRGSMSYTKLSGL 708
CC 181 TLLAVALIAPFMGAVFSGITVYCVCDHRKDVAVQKKKELTSSRGSMSYTKLSGL 240
CC 709 FGDTSKDPKPRPAITPLMNGKLTAPNTAAMLTKAQDHLDTALPTPESTPTLQQR 768
CC 241 FGDTSKDPKPRPAITPLMNGKLTAPNTAAMLTKAQDHLDTALPTPESTPTLQQR 300
CC 769 KPSRSGRENNQNTINACTKMPMGSPVITDPLPLASPHISVAVVLPITQGYGHE 828
CC 301 KPSRSGRENNQNTINACTKMPMGSPVITDPLPLASPHISVAVVLPITQGYGHE 360
CC 829 YVDQPKSEVAAQALDQALTEYKTIKELSSKSNHGVNENLDSLPFVPOREASL 888
CC 361 YVDQPKSEVAAQALDQALTEYKTIKELSSKSNHGVNENLDSLPFVPOREASL 420
CC 889 GPPGASISQGLSKLEMHSSSYGVYKRSYPTNSLTRSHQATTLKNNNTSSNSHLS 948
CC 421 GPPGASISQGLSKLEMHSSSYGVYKRSYPTNSLTRSHQATTLKNNNTSSNSHLS 480
CC 949 RNQSGFRGNNPPAPQVRVDSIQVHSGQPGAVTVSRQSLNAYNSLTRSGIKRTPSLKP 1008
CC 481 RNQSGFRGNNPPAPQVRVDSIQVHSGQPGAVTVSRQSLNAYNSLTRSGIKRTPSLKP 540
CC 1009 DVEPKSPAPLSTSMKPNDACT 1030
CC 541 DVEPKSPAPLSTSMKPNDACT 562
CC DB

RESULT 11
ABG04066
ID ABG04066 standard; Protein; 863 AA.
XX
AC ABG04066;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #4057.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX
PN WO200175067-A2.

XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001MO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
XX
XX (HYSO-) HYSO INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI: 2001-639362/73.
XX N-PSDB; AAS68253.
XX
XX New isolated polynucleotide and encoded polypeptides useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity -

Claim 20; SEQ ID No 34425; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic amino acid sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at [ftp://wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).

Sequence 863 AA:

Query Match 53.0%; Score 2891; DB 22; Length 863;
Best Local Similarity 80.9%; Pred. No. 5.9e-242;
Matches 583; Conservative 28; Mismatches 92; Indels 18; Gaps 8;

CC 321 STPGSAVAYMDLIDASVFTGRFKQKSPDSTWTPDERVYKPRPGCCAGSSSLERYAT 380
CC 150 STPGSAVAYMDLIDASVFTGRFKQKSPDSTWTPDERVYKPRPGCCAGSSSLERYAT 209
CC 381 SNEPPDDTLNFKTHPLMDEAVPSIFNRPWFLRTWRYRLTKIANDTAAGPQNTTVFL 440
CC 210 SNEPPDDTLNFKTHPLMDEAVPSIFNRPWFLRTWRYRLTKIANDTAAGPQNTTVFL 269
CC 441 GSEKGIILKFLARLNGSGFLNLSFLEMSVYNSKCSYDGVEDKRI-MGMQDRASSSL 499
CC 270 GSEKGIILKFLARLNGSGFLNLSFLEMSVYNSK-KWSTYKAPRVITILNPGASFCFI 328
CC 500 YVAFSTCV---IKYPLGRG--ERHGKCKKCTCIASRDPYC--GMIKEGAGCSHLSPLNSR 550
CC 329 TIRETYCARKKRIWCPRYCTIOXH-FCPCHLISGKETLCRVTGKMKVADRDESLPYAA 387
CC 551 LTFEODI-ERGNTDGLDCHNSFVALNGHSSSLPSTTTSDTAQEGYSGRGMMDKHL 609
CC 388 MLLAQOMARCKEELGITLAIHKRAVGNRTK---TGPASPRSSPCPLGLKWQTL 442
CC 610 LQSPDSTPLGAVSSHHNODKKGYTRRSYTKGHDQVPTLLAVALIAPFMGAVFSGIT 669
CC 443 FPRRLRMFGGRRRSQLEAQRVIRESYLKGHDQVPTLLAVALIAPFMGAVFSGIT 502
CC DB

QY VYVCDHRRKDVAVVORKEKELTHSRGSMGVTLKSGLPDQSKDEKPEALITPLMHN 729
 Db 503 VYVCDHRRKDVAVVORKEKELTHSRGSMGVTLKSGLPDQSKDEKPEALITPLMHN 562
 QY 730 GKLATPGNTAKMLIKADQHHLDLTLPTPESTPTLQCKRKSRSRGRERERNQNLINACTK 789
 Db 563 GKLATPGNTAKMLIKADQHHLDLTLPTPESTPTLQCKRKSRSRGRERERNQNLINACTK 622
 QY 790 DMPKMGSPVITPDLPLRASPSHISPVVLPITQCGYQHEVVDQPKMSVAVQALMLEDQAAT 849
 Db 623 DMPKMGSPVITPDLPLRASPSHISPVVLPITQCGYQHEVVDQPKMSVAVQALMLEDQAAT 682
 QY 850 LEKTKIKELSKSNHGVNLVENDSLPKVPOREASIGPPGASLSQTLGSKRLEMHS 909
 Db 683 LEKTKIKELSKSNHGVNLVENDSLPKVPOREASIGPPGASLSQTLGSKRLEMHS 742
 QY 910 SSGYVDYKRSYPTNSLTRSHQATTLKRNNTSSNSSHSRNOSFGRGDNPAPAPQVDSI 969
 Db 743 SSGYVDYKRSYPTNSLTRSHQATTLKRNNTSSNSSHSRNOSFGRGDNPAPAPQVDSI 802
 QY 970 QVHSSQPSQAVTVSRQPSLNAYNSLTRGLKRTPSLXKQDVPKPSFAPLSTSMKPNDAC 1029
 Db 803 QVHSSQPSQAVTVSRQPSLNAYNSLTRGLKRTPSLXKQDVPKPSFAPLSTSMKPNDAC 862
 QY 1030 T 1030
 Db 863 T 863

RESULT 12

ABG04620 standard; Protein; 962 AA.

ID ABG04620 standard; Protein; 962 AA.
 XX ABG04620;
 AC 13-FEB-2002 (first entry)
 DT 13-FEB-2002 (first entry)
 XX Novel human diagnostic protein #4611.
 DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW Food supplement; medical imaging; diagnostic; genetic disorder.
 XX Homo sapiens.
 OS WO200175067-A2.
 XX 11-OCT-2001.
 PD 30-MAR-2001; 2001WO-US08631.
 PE 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX (HYSE-) HYSEQ INC.
 XX Drmanac RT, Liu C, Tang YT;
 PI WPI; 2001-639362/73.
 DR N-PSDB; AAS68807.
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 Claim 20; SEQ ID No 34979; 103bp; English.
 The invention relates to isolated polynucleotide (I) and
 polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags

CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 962 AA;

Query Match 52.8%; Score 2879.5; DB 22; Length 962;

Best Local Similarity 70.3%; Pred. No. 7.1e-241; Mismatches 122; Indels 101; Gaps 12;

Matches 599; Conservative 30; Mismatches 122; Indels 101; Gaps 12;

QY 147 YKMDT-LEPFDFSGMARCPYDAKHAVALFA-----DGLYSATVTPFLAD 194
 Db 95 YEPDTWLGKSSMTLVFTHPFSAALTHSALARSDDTSLSPDGKYSATVTPFLAD 154
 QY 195 AVIYRSLGSPPLRTVYKHSKWLKEPYFQAVDYGDYIFFPRELAVEYNTMGKVPFPRV 254
 Db 155 AVIYRSLGSPPLRTVYKHSKWLKEPYFQAVDYGDYIFFPRELAVEYNTMGKVPFPRV 214
 QY 255 AOVCKNDMGSGORVLEKQWTSFLKARLNCVSDSHFENILQAVTVIRINGRDVYLAT 314
 Db 215 AOVCKNDMGSGORVLEKQWTSFLKARLNCVSDSHFENILQAVTVIRINGRDVYLAT 274
 QY 315 FSTPYNSIPGSAVCAVMDLIDIASVFTGRFEKQSPSTWTPVDEVPKPPRCCGAGSS 374
 Db 275 FSTPYNSIPGSAVCAVMDLIDIASVFTGRFEKQSPSTWTPVDEVPKPPRCCGAGSS 334
 QY 375 LERVATSNFPPDPTLNFITKTHPLMDEAVPSIFRPPFLRMVYRRLTKLAVDPAAGPYON 434
 Db 335 LERVATSNFPPDPTLNFITKTHPLMDEAVPSIFRPPFLRMVYRRLTKLAVDPAAGPYON 394
 QY 435 HTVAVPLGSEKGIILKEFARIGNSGFINDSLFLEMSVYNSEKSYDGEDKRIHQOLDR 494
 Db 395 HTVAVPLGSEKGIILKEFARIGNSGFINDSLFLEMSVYNSEKSYDGEDKRIHQOLDR 454
 QY 495 ASSSLVYAFSTCYIKVPLGRCEHKGCKKTCIASRDYCGWIKEGACSHLSPNSRLT-- 552
 Db 455 ASSSLVYAFSTCYIKVPLGRCEHKGCKKTCIASRDYCGWIKEGACSHLSPNSRLT-- 514
 QY 553 FPODIERGNTDGLGDC-----HNSFVALNGHSS----- 581
 Db 515 FVVPVVTGST--LHRCILLTSLMRQTHPSY-----HSSSEGRPVKGERGKCTLNHFEL 567
 QY 582 --LPSITTSDDTAQEGYESRQGV-----LDWKHLDSPDSTDLGAVSSHNHODKCVI 634
 Db 568 GYWLPEVSYTLRVAIVGVGLKRLQTLFPRRLRM-----PQGRKKKSQLEAQKVI 616
 QY 635 RESYLKGGHDLVPTTLAIAVILAIFNMGAVPSGITYVCVCHRRKDVAVVORKEKELTHS 694
 Db 617 RESYLKGGHDLVPTTLAIAVILAIFNMGAVPSGITYVCVCHRRKDVAVVORKEKELTHS 676
 QY 695 RRGMSAVTKLGSGLFDQSKDPPEALITPLMNGKLATPGNTAKMLIKADQHHLDLTLA 754
 Db 677 RRGMSAVTKLGSGLFDQSKDPPEALITPLMNGKLATPGNTAKMLIKADQHHLDLTLA 736
 QY 755 LPTPESTPTLQCKRKSRSRGRERERNQNLINACTQDMPKMSVAVQALMLEDQAATPLRASPSHS 814
 Db 737 LPTPESTPTLQCKRKSRSRGRERERNQNLINACTQDMPKMSVAVQALMLEDQAATPLRASPSHS 796
 QY 815 VVVLPTQCGYQHEVVDQPKMSVAVQALMLEDQAATLEKTKIKELSKSNHGVNLVENDL 874

Db 797 WMSCPSSRSATSMSTWTSPPKXAFPMRWMTTRPMPHSIRPSRNIISAARVPTMGXTLMRTW 856
 QY 875 DSLPPK-----VPGREASLGPFGAS-----LSQTGSKRLEMHSSSYGVDRSY 920
 Db 857 TACPPKHSGRPPWPVREPCLRPVXASGKCTTPLPLGLTIR----- 899
 QY 921 PTNSLTRSHQAT 932
 Db 900 GATPRTRSRKAT 911
 Db 900 GATPRTRSRKAT 911
 RESULT 13
 AAB92688
 ID AAB92688 standard; Protein: 507 AA.
 AC AAB92688;
 XX 26-JUN-2001 (first entry)
 DT 26-JUN-2001 (first entry)
 DE Human protein sequence SEQ ID NO:11073.
 XX Human protein sequence SEQ ID NO:11073.
 DE Human protein sequence SEQ ID NO:11073.
 XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 KW Homo sapiens.
 OS Homo sapiens.
 PN EPI074617-A2.
 XX EPI074617-A2.
 PD 07-FEB-2001.
 PF 28-JUL-2000; 2000EP-0116126.
 PR 29-JUL-1999; 99JP-0248036.
 PR 27-APR-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX 09-JUN-2000; 2000JP-0241899.
 PA (HELI-) HELIX RES INST.
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI; 2001-318749/34.
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 PS Claim 8; SEQ ID 11073; 2537bp + CD ROM; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialized methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.

XX Sequence 507 AA;
 SQ Query Match 48.1%; Score 2621.5; DB 22; Length 507;
 Best Local Similarity 90.2%; Pred. No. 7e-219; Indels 55; Gaps 1;
 Matches 507; Conservative 0; Mismatches 0
 QY 469 MSYVNSFKSGYDVEPKRIWGMQLDRASSSLYVAFSTCVIKVPLGRCEHKKCTCIAS 528
 Db 1 MSYVNSFKSGYDVEPKRIWGMQLDRASSSLYVAFSTCVIKVPLGRCEHKKCTCIAS 60
 QY 529 RDPYCGWIKKEGACSHLSPNSRLTFEODIERGNTDGLGDCNHFVALNGHSSLLPSTTT 588
 Db 61 RDPYCGWIKKEGACSHLSPNSRLTFEODIERGNTDGLGDCNHFVALNGHSSLLPSTTT 108
 QY 589 SDSTAEGYSGRGMLDMKHLDSBDSSTDEPGANSHNHQDKGVIRESYLKGHDQVYV 648
 Db 109 -----GVIRESYLKGHDQVYV 125
 QY 649 TLLAVIILAFVWGAVFSGITVYCVCDHRRKDVAVVQKKEKELTHSRGSMSSVTKLSGL 708
 Db 126 TLLAVIILAFVWGAVFSGITVYCVCDHRRKDVAVVQKKEKELTHSRGSMSSVTKLSGL 185
 QY 709 FGDTSKDPKPEALITPLMHNGKLTAPGNTAKMLIKADQHLDTALPTPESTPTLOQR 768
 Db 186 FGDTSKDPKPEALITPLMHNGKLTAPGNTAKMLIKADQHLDTALPTPESTPTLOQR 245
 QY 769 KPSRGRWRERNONLINACTKMPMGSPVIFPTDPLASPSHIFSVVVLPIIOGYQHE 828
 Db 246 KPSRGRWRERNONLINACTKMPMGSPVIFPTDPLASPSHIFSVVVLPIIOGYQHE 305
 QY 829 YVDQPKMEVAQMALEDQAATLEYKTIKEHLSKSPNNGVNLVENLDSLPRKVPQRESL 888
 Db 306 YVDQPKMEVAQMALEDQAATLEYKTIKEHLSKSPNNGVNLVENLDSLPRKVPQRESL 365
 QY 889 GPPGASLSQGLSKRLEMHSSSYGVDRKSYPTNSLTRSHQATTLKRNTTSSNSHLS 948
 Db 366 GPPGASLSQGLSKRLEMHSSSYGVDRKSYPTNSLTRSHQATTLKRNTTSSNSHLS 425
 QY 949 RNQSGRGRGNPPAPAPRVDSIOVHSSQPSGCAVATSRQPSLNAVNSLTRSGLKRTPSLKP 1008
 Db 426 RNQSGRGRGNPPAPAPRVDSIOVHSSQPSGCAVATSRQPSLNAVNSLTRSGLKRTPSLKP 485
 QY 1009 DVPKPSFAPLSTSMKENDACT 1030
 Db 486 DVPKPSFAPLSTSMKENDACT 507
 RESULT 14
 ID AAG63213
 ID AAG63213 standard; Protein: 1086 AA.
 AC AAG63213;
 DT 01-OCT-2001 (first entry)
 XX
 DE Amino acid sequence of a human semaphorin-like polypeptide.
 XX
 KW Expressed sequence tag; EST; semaphorin-like protein; neuronal growth;
 KW spinal cord damage; neurodegenerative disease; genetic neuronal defect;
 KW immunological disorder; lymphocyte dysfunction; viral infection; cancer.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..16
 FT /note= "signal peptide"
 FT 671..699
 FT Domain /note= "transmembrane domain"
 XX
 EN WC200153466-A1.
 XX
 PD 26-JUN-2001.

10-DEC-2001; 2001WO-US48369.
 08-DEC-2000; 2000US-254329P.
 14-DEC-2000; 2000US-255648P.
 15-MAY-2001; 2001US-291037P.
 08-JUN-2001; 2001US-297173P.
 08-JUN-2001; 2001US-309258P.
 29-AUG-2001; 2001US-315639P.
 01-OCT-2001; 2001US-326393P.
 (CURA-) CURAGEN CORP.

Alsobrook JP, Anderson DW, Burgess CE, Boldog FL, Casman SJ,
 Coleman SD, Edinger SR, Ellerman K, Gerlach V, Gorman L, Grosse WM,
 Guo X, Herrmann JL, Kekuda R, Lepley DW, Li L, MacDougall JR,
 Millet I, Pena CA, Peyman JA, Rastelli L, Rieger DK, Shinkets RA,
 Smithson G, Speyer KA, Stone DJ, Tchernev VT, Vernet CAM, Voss EZ,
 Zernhusen BD, Zhong H, Zhong M;
 WPI; 2002-643486/69.
 N-PSDB; ABS64384.

New NOVX polypeptides and polynucleotides useful for treating or
 preventing e.g. neurodegenerative diseases, neurological disorders,
 cardiovascular diseases, muscular diseases and disorders, or
 immunological diseases

Claim 1; Page 58; 29pp; English.

The present invention relates to new NOVX polypeptides. The polypeptides,
 polynucleotides and antibodies are useful in the manufacture of a
 medicament for treating or preventing neurodegenerative diseases (e.g.
 Alzheimer's disease, Parkinson's disease, or Huntington's disease),
 neurological disorders (e.g. anxiety, schizophrenia, manic depression or
 mental retardation), cardiovascular disease (e.g. acute heart failure,
 angina pectoris or myocardial infarction), muscular diseases and
 disorders, retinal diseases (including those involving photoreception,
 deafness and keratinisation disorders), cancer (e.g. ovarian cancer or
 melanoma), immunological disorders, inflammatory and immune diseases,
 bacterial, fungal, protozoal and viral infections, and reproductive
 system disorders. The proteins of the invention may be used to screen
 drugs or compounds that modulate the NOVX protein activity or expression,
 as well as to treat disorders characterised by insufficient or excessive
 production of NOVX protein or protein forms that have decreased or
 aberrant activity compared to NOVX wild type protein, such as diabetes,
 obesity, metabolic disturbances associated with chronic diseases and
 wasting disorders associated with chronic diseases and various cancers,
 infectious diseases and various dyslipidaemias. The nucleic acid
 sequences of the invention may be used in chromosome mapping,
 identifying an individual from minute biological samples (tissue typing),
 and in forensic identification of a biological sample. The present
 amino acid sequence represents a NOVX protein of the invention.

Sequence 1088 AA;

Query Match 43.1%; Score 2349.5; DB 23; Length 1088;
 Best Local Similarity 45.4%; Pred. No. 1.2e-194;
 Matches 512; Conservative 188; Mismatches 309; Indels 149; Gaps 31;

1 MSEALLLYFTLL--HFAAGFPEDESEPIST--SHGNTTKQYPPVVGKPKRNTTQRH 55
 1 MRFVLLCAYILLMLWSQLRAVSPEDDEPLNTVDYHCKSSRQPPVRG--RFSGNSQC-HR 58
 56 LDIQIMIMINGTITARADHIYVDIDTSHTEIYSCKLTKWRSQADVDTGKMGKHKD 115
 59 LDPQIMIKIRDTIYIAGDQVYVNLNEMKPTVIMQCKLTWMSRQDRENCKMGKHKD 118
 116 ECHNFIKVLKKNDDALFVGTNAFNPSCNRYKMDLLEPFGEFSGMARCPIYAKANVA 175
 119 ECHNFIKVPFRDDEWVFGTNAFNPSCNRYKMDLLEPFGEFSGMARCPIYAKANVA 178
 176 LPADGKLYSATVTDPLAIDAVIYRSIGSEPTLRTVKHDSKMLKEPYVQAVDYGDIYFF 235

179 LPADGKLYSATVTDPLAIDAVIYRSIGSEPTLRTVKHDSKMLKEPYVQAVDYGDIYFF 238
 236 FREIAVEYNTMGKVFPRVAVQVCKNDMGSGQVLEKQMTSFLKARLNCVPGDSHFYENI 295
 239 FREIAVEHNNLKGAVYRVAIKCNNDMGSGQVLEKQMTSFLKARLNCVPGDSHFYEDV 298
 296 LOAVTDVIRINGRDVAVLATFSTPNYSIPGSAVCAVMDLIVAFVTFGRFKEQKSPDSTWTP 355
 299 LOSTIDIIQINGIPLVAVVFTTQUNSIIPGSAVCAVSMIDIEKVFGRFKEQKTPDSVWTA 358
 356 VPDERVPRPRPGCCAGSSSLERVATSNFPDPTLNFITKTHPLMDAVPSIFNRPMFLTM 415
 359 VPEDKVPFRPRGCCAGKGLAEVYKTSIDFPDETUSFISHPLMDSAVPIADPEWFTKTR 418
 416 VRYRLTKAVDTAAGPYQNHVVFELSGEKIITKELARLNGSGF--LNDSLFLEMSVYNS 474
 419 VRYRLTAIVDSHAPYQNYVTFVSGSAGWLVKVLAK--TSPFLNDSVLLLEAVNH 476
 475 EKSVDGVDEKRIWGMOLDRASSLYVAFSTCVIVPLRCERHCKKCTCIASDPYCG 534
 477 AKCSAENEDKRVISLQDKDHALVAFSSCIIRIPSRCEYSGCKKSCIASDPPYCG 536
 535 WKEGACGCHLSPNRRLT-----FEODIERGNTDGLGDCCHNSFVALNGHS 580
 537 WLSQ--GSGRVTPGMILLTEDEFAFNHSAEGYEDTERGNTAHLDCH----- 584
 581 SLPLSTTSDSTAQSEYSESG-----MMD--WKHL----- 610
 585 EILPTSTPD-----YKIFGQPTSMVEYSSSVTTMASIPETPKVIDTWPRKLTSSRK 638
 611 ----DSP--DSTDPILGAVSHNHQDKGVIRESYKGDQVPLVTLIAVLAIVMGA 663
 639 FVQDDPNNSDFDELSTG-----PKGVMEVQGESQVNMVNLITCVAAAFVLA 692
 664 VFSGITVYCVCD--HRRKDAVAVQKRELTHSRGSMSSVTKLSGF--DTSKDPK 718
 692 FIAGVAVYCRADMEVFNKRI--HKDAESAQSTDSGSAKLNGLFDSVKEVQONIDS 749
 719 PEALIPPLMNGKLAATGNTAKMLIKADQHLDLALPTPESTPTLQOKRPSRGEWE 778
 750 PKLYSNLTSRKELPPNGDTKSMVMDHROGPBELAALPTPESTVHLQKTLQAMKSHSEK 809
 779 RNQMLINACTKMPWMSPIPTDLPFASPSHIPSVVLPITQOQYQHEV----- 829
 810 AHGH--GASRKETQFPSSPPHSPV--SHGILPSAIVPMATHDVNTSFSNNAKAE 865
 830 ----VQPR--KMSVAYQALBEDQATLEKTIKELHSSKSPN-----HGVNL 870
 866 KKLQNIHPLTKSSSKDRHRRSVDSRNTL--NDLLKHLNDNSNPKALMGDIQMAHQVLM 923
 871 VENIDSL--PPKVPREASIGPGASISQGLSKRLEMHSSSY--GVDPKSYTPNS 924
 924 LDPWGSMSVPPKVPKPNRASLYSPSTLPRNSPTKRVAVPTTPGVPMTSLRORGYHNS 983
 925 LTRSHOATTLKRNNTNSNSHLSRNOFGAGD--NPPAPQRYDSIOVHSSQPSQCAVTV 983
 984 SGR--HSIAMPK--NLNSPNGVLLSKROPENMGKGYPTTGAKVDTIQ-----GTPYSV 1034
 984 SROPELANYNSTL-----RSGIKRTPSLKDPVPKPSFAPLSTMKP 1025
 1035 HLPQSLRQSSYTSNGTLPTGLKRTPLSLKRPVPPKPSFVPTPSVPR 1082

Search completed: October 23, 2003, 17:10:34
 Job time : 56 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: October 23, 2003, 17:09:07 ; Search time 19 Seconds
(without alignments)
2293.694 Million cell updates/sec

Title: US-09-856-681-2

Perfect score: 5450

Sequence: 1 MRSEALLLYFTLHPAGAGF.....PKPSPAPLSTSMKPNDACT 1030

Scoring table: BLOSUM62
Gapop 10.0 , Gapex 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2013.5	36.9	888	4	US-09-077-940A-4
2	1597.5	36.6	887	4	US-09-077-940A-2
3	1519.5	27.9	930	4	US-09-254-594-6
4	1428.5	26.2	929	4	US-08-121-713D-58
5	969.5	17.8	730	1	US-08-121-713D-58
6	969.5	17.8	730	1	US-08-835-268-58
7	969.5	17.8	730	2	US-09-060-692-58
8	969.5	17.8	730	3	US-08-833-391-58
9	969.5	17.8	730	4	US-09-060-610-58
10	969.5	17.8	730	5	PCT-US94-10151A-58
11	880	16.1	650	1	US-08-121-713D-60
12	880	16.1	650	1	US-08-835-268-60
13	880	16.1	650	2	US-09-060-692-60
14	880	16.1	650	3	US-08-833-391-60
15	880	16.1	650	4	US-09-060-610-60
16	880	16.1	650	5	PCT-US94-10151A-60
17	852.5	15.6	655	4	US-08-556-422A-3
18	852.5	15.6	655	1	US-08-121-713D-54
19	852.5	15.6	655	1	US-08-835-268-54
20	852.5	15.6	655	2	US-09-060-692-54
21	852.5	15.6	655	3	US-08-833-391-54
22	852.5	15.6	655	4	US-09-060-610-54
23	852.5	15.6	655	5	PCT-US94-10151A-54
24	845	15.5	712	1	US-08-121-713D-64
25	845	15.5	712	1	US-08-835-268-64
26	845	15.5	712	2	US-09-060-692-64
27	845	15.5	712	3	US-08-833-391-64

28	845	15.5	712	4	US-09-060-610-64	Sequence 64, Appl
29	845	15.5	712	5	PCT-US94-10151A-64	Sequence 64, Appl
30	779.5	14.3	724	1	US-08-121-713D-62	Sequence 62, Appl
31	779.5	14.3	724	1	US-08-835-268-62	Sequence 62, Appl
32	779.5	14.3	724	2	US-09-060-692-62	Sequence 62, Appl
33	779.5	14.3	724	3	US-08-833-391-62	Sequence 62, Appl
34	779.5	14.3	724	4	US-09-060-610-62	Sequence 62, Appl
35	779.5	14.3	724	5	PCT-US94-10151A-62	Sequence 62, Appl
36	748	13.7	775	4	US-09-308-179B-1	Sequence 1, Appl
37	736.5	13.5	862	4	US-08-556-422A-2	Sequence 2, Appl
38	654.5	12.0	607	4	US-08-556-422A-4	Sequence 4, Appl
39	586	10.8	477	1	US-08-136-922-2	Sequence 2, Appl
40	539	9.9	425	4	US-08-556-422A-7	Sequence 7, Appl
41	462	8.5	295	4	US-08-556-422A-6	Sequence 6, Appl
42	437	8.0	634	3	US-09-041-236-2	Sequence 2, Appl
43	437	8.0	634	4	US-09-771-467C-2	Sequence 2, Appl
44	437	8.0	666	3	US-09-240-410-2	Sequence 2, Appl
45	356.5	6.5	606	3	US-09-041-236-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-09-077-940A-4
Sequence 4, Application US/09077940A
Patent No. 6576441
GENERAL INFORMATION:
APPLICANT: KIMURA, Toru et al.
TITLE OF INVENTION: NOVEL SEMAPHORIN 2 AND GENE ENCODING THE SAME
FILE REFERENCE: 0020-4426P
CURRENT APPLICATION NUMBER: US/09/077,940A
CURRENT FILING DATE: 1998-06-05
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 888
TYPE: PRT
ORGANISM: Homo sapiens
US-09-077-940A-4

Query Match 36.9%; Score 2013.5; DB 4; Length 888;
Best Local Similarity 45.3%; Pred. No. 1.7e-182;
Matches 424; Conservative 129; Mismatches 241; Indels 141; Gaps 20;

QY	5	ALLLYFTLHPAGAGFEDSEPTISISGNVYKQVPGVGHKGGNTTOR--HRLDQIM 62
DB	12	ALLLLLLLGGAGHLPEDPPPLSVAPRDYLNHPVFGSGPKLTPAEGADLNIQVYL 71
QY	63	IMNGTLVYIARDHIYTVDDIDTSHTTEIYCSKLTWKSQADVDTCRMKGKHKDECHNFIK 122
DB	72	RVNRTLFIQBDNLRYRELEFPPTSTELRYQKLTMRSPSDINVCRMKGKGEGBRNFVK 131
QY	123	VLLKKNDALFVCGTNAENPSCRYKMDLEPPFGDESGMARCPYDAKHANVALPADQKL 182
DB	132	VLLLRDESTLVCSSNMFPCANYSIDTLQPVGDNISGMARCPYDPGHANVALFSDQML 191
QY	183	YSAVYTTDLADDAVYIRSLGSPILRTYKHSKMYKEFYQAVDYGVIYFEFEIYAVE 242
DB	192	FTATVTDLADDAVYIRSLGSPILRTYKHSKMYKEFYQAVDYGVIYFEFEIYAVE 251
QY	243	YNTMGKVVFPVAVQCKDKMGSGORVLEKQMTSFLKARLNSVDPGSHFFNYIIQAVTV 302
DB	252	FNYLEKVVSVKARVCKDKDVGGSPVLEKQMTSFLKARLNSVDPGSHFFNYIIQAVTV 311
QY	303	IRINGRDVVLATFTFPYNSIPGSAVCAVMDIASVFTGRFKRQKSPDSTWTFPVDEKVP 362
DB	312	VSLGRPVVLAVFTSPNSIPGSAVCAVMDIASVFTGRFKRQKSPDSTWTFPVDEKVP 371
QY	363	KRPBCCAGSSLEKAYATSNFEDDTLNFTCHPLMDAVSIFNRPFLRTMVFYRLTK 422
DB	372	RPRBCCAPGM--QYNASSALPDILNFVKTHPLMDAVFSLGHPMLRLTKMHQHLTR 429


```

QY 423 IAVDTAAGYQNHVTVFLGSEKGIILKEFLAR--IGNSGFLNDSLLEEMSVYNSEKCYD 480
DB 430 VAVDVAGAGPWGQTVTVFLGSEAGTVLKFVLRPNASTSGTSLVLEEFETRRPDRCGRP 489
QY 481 GVED--KRTMGQDLBRASSSLVVAFTCVIKVPLRCGRHCKCKTKCIASRPYCGWIK 538
DB 490 GGEETGQRLSLLELDASGGLLAAPRCVAVPVAAQQYSGCMKNCIGSQDPYCGWAD 549
QY 539 GGACSHLSPNSRLTFEODIERGNTDGLGCHNSFVALNGHSSSLPSTTSDTAQEGYE 598
DB 550 -GSCIFLSTGCTAAAFQDVSGASTGSLGDC----- 578
QY 599 SSGKMLDWKHLLDSPDSTPLGAVSSHNDQKKGVIRBSYLGKHDQLYPTLLAVALIA 658
DB 579 -----TGILRASLSDDRAGLVSVMILVTSVA 606
QY 659 FVMGAVFSGITV-YVCCHRRKDVAVVORKEKE--LTHSRGSMSSVTLSGLFGDTQK 715
DB 607 FVVGAVVSGFVGMFGLRERRELA--RRKDEKAILAHGAGEAVLSVSL---GERRAQ 660
QY 716 DP-----KPEAILTPLMNGKLTATPGNTAK-MLIKADQHLDTALTPTPEST 761
DB 661 GGGGCGGGGGGAGVPEALIMQNG-----WAKATLQGGPHDLDSGLPTPEQT 713
QY 762 PTLQCKRP-----SRGSEMERONLINAOTKMPMGSPVPTDLP---LRASPEH 811
DB 714 P-LPQKRLPTPHPHALGPRAMDH-----GHPLPASAASSLLLLAPAR 757
QY 812 ISVVVLPITQGYGHEVVDQPKMSEVAAQMALEDQAATLEYKTIKELHASKSPNHQVNLV 871
DB 758 ABQPPAPBEPPTDGLIYARPERASHGDFPLTPHASPBRKRV-----SAPTGLDPA 811
QY 872 ENLDSLP-PKVPQREASL-----GPPGASLSQT 898
DB 812 SAADGLPRPMSPPPTGSLRRPLGPHAPPAATLRT 846

```

RESULT 2
 US-09-077-940A-2
 ; Sequence 2, Application US/09077940A
 ; Patent No. 6576441
 ; GENERAL INFORMATION:
 ; APPLICANT: KIMURA, Toru et al.
 ; TITLE OF INVENTION: NOVEL SEMAPHORIN 2 AND GENE ENCODING THE SAME
 ; FILE REFERENCE: 0020-4426P
 ; CURRENT APPLICATION NUMBER: US/09/077,940A
 ; CURRENT FILING DATE: 1998-06-05
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 2:
 ; LENGTH: 887
 ; TYPE: PRT
 ; ORGANISM: Rattus norvegicus
 US-09-077-940A-2

```

Query Match 36.6%; Score 1997; DB 4; Length 887;
Best Local Similarity 42.7%; Pred. No. 6,3e-181;
Matches 433; Conservative 132; Mismatches 260; Indels 190; Gaps 25;

```

QY 2 RSEALLYFTLLHFAAGFPEDESEPTISHSNNTKQVAVVGHKPGKNTQR--HRLDIQ 59
 DB 10 RPALELLLLLRVTHGLFPEDEPPLVAAPRDYLSHFVVGSGPRRLTAEGADLNIG 69
 QY 60 MIMINGSLYIARADHIYVDIDTSHTETIYCSKSLTKWKSROADVDTCKMKGKHDECHN 119
 DB 70 RVLRVNRTFLIGRDNLYVELLEPSTSTELRYQKLTWRNPISDIDVCRMKGQEGECRN 129
 QY 120 FIKVLKKNDALFVCGTNAFNSCRANYKMDLEPFDESSGACGYDAKXNVALLAD 179
 DB 130 FVKVLLRDSSTLFVCCSNFNPICANYSMDTQLQGDNTSGWRCYDPRKANVALFSD 189
 QY 180 GKLYSATVTDPLADAVIYSLGSPFLRTVKHDSKWLKEPFYQAVDGYIYFFPREI 239

```

DB 190 GMLPTATVTDPLADAVIYSLGDRPFLRTVKHDSKWKKEYFVHAWEWSHVYEFFREI 249
QY 240 AVEYNTMGKVVFPRAVQYCKNDMGGSQYLEKQMTSEFKARLNCVSPEDSHFENILQAV 299
DB 250 AMEFVYLEKVVSVARVAKCKNDVGSFVLEKQMTSEFLKARLNCVSPEDSHFENILQAV 309
QY 300 TDVIRINRDVVLATFSPNYSIGSANCAVMDMDIASVFGREFECKSPSTMTVPVDE 359
DB 310 TGVVSLGRPVLIAVFTSPNSIPGSVACADIMOVAAVFESRFEQSPSIIWTFVPEED 369
QY 360 RVKPRPGCCAGSSSLEBYATSNNEPDDTLNFIKTHPLMDEAVPSIFNRPMFLRTMVRYR 419
DB 370 QVPRRPFGCCAPAGM-QYNASNAIPDEILNFVKTHPLMDAVPGLSGSPWIVRLIHHQ 427
QY 420 LTKIADVTAAGYQNHVTVFLGSEKGIILKEFLAR--IGNSGFLNDSLLEEMSVYNSEK 477
DB 428 LTRVADVAGAGPWGQTVTVFLGSEAGTVLKFVLRPNASTSGTSLVLEEFETRRPDR 487
QY 478 ---SYDVEDKRIKMGQDLBRASSSLVVAFTCVIKVPLRCGRHCKCKTKCIASRPYCG 534
DB 488 GRSSAGAGWQRLSLLELDASGGLLAAPRCVAVPVAAQQYSGCMKNCIGSQDPYCG 547
QY 535 WKEGACSHLSPNSRLTFEODIERGNTDGLGCHNSFVALNGHSSSLPSTTSDTAQ 594
DB 548 WAPD-GSCIFLRPGTSATFEQDVSGASTGSLGDC----- 580
QY 595 EGYESRGKMLDWKHLLDSPDSTPLGAVSSHNDQKKGVIRBSYLGKHDQLYPTLLAIA 654
DB 581 -----TGILRASLSDDRAGLVSVMILVTS 604
QY 655 VILAIVGAVFSGITV-YVCCHRRKDVAVVORKEKE--LTHSRGSMSSVTKL----- 705
DB 605 SYAAIVVAGVAVSGFVGMFGLRERRELA--RRKDEKAILAHGAGEAVLSVRLGRRGT 662
QY 706 --SGFQDTQSKDPREAILTPLMNGKLTATPGNTAKMLIKADQHLDTALTPTPESTPT 763
DB 663 GTGGGAGAGGGGAGPEALIMQNG-----TKAALLHGGPHDLDSGLPTPEQTP 715
QY 764 LQCKRPKPR-----GSEMERONLINAOTKMPMGSPVPTDLP---DMPMGSPVPTD 802
DB 716 LPQKRLPTTHPHALGPRAMDHSHALLSASASTLLLIATRAPQPP-----VTEBG 770
QY 803 -----LPLRASPSPHPSVVVLPITQGYGHEVVDQPKMSEVAAQMALEDQAATLEYKTIK 856
DB 771 PESRLCAPSGASHGDEPLTP-----HASPDRRVVSAPTGLDSSVG----- 816
QY 857 EHLSSKSPNHGVNLVENLDSLP-PKVPQREASL-----GPPGASLSQTKSLKLEMHHS 909
DB 817 -----DDLPGPMSPPATSSLRPRGPHGPPPAALRKT-----HT 849
QY 910 SSGYVDYKRSYPTNLSLTRHQATTLKRNNTSNGSHSRNOSFGGD-NPPAP 963
DB 850 FMSG---EAPGGRPRRHA-----PADSTHL--LPCGIGEXTAPPVP 887

```

```

RESULT 3
US-09-254-594-6
; Sequence 6, Application US/09254594
; Patent No. 6566094
; GENERAL INFORMATION:
; APPLICANT: KIMURA, Toru
; APPLICANT: KIKUCHI, Kaoru
; TITLE OF INVENTION: NOVEL SEMAPHORIN GENE: SEMAPHORIN Y
; FILE REFERENCE: 0020-4527P
; CURRENT APPLICATION NUMBER: US/09/254,594
; CURRENT FILING DATE: 1999-05-11
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:

```

NAME/KEY: misc feature
 LOCATION: ()..()
 OTHER INFORMATION: Tissue Type: Child Brain
 NAME/KEY: misc feature
 LOCATION: ()..()
 OTHER INFORMATION: Identification Method: P for resulting peptide
 us-09-254-594-6

Query Match 27.9%; Score 1519.5; DB 4; Length 930;
 Best Local Similarity 35.94; Pred. No. 2,4e-135;
 Matches 381; Conservative 135; Mismatches 356; Indels 189; Gaps 32;

6 LLLYFTLLHPAGAFPEDESEPISSHGNYTKQYVFNVGHKRGKNTQRRHLDIQMIMIN 65
 13 LLLLSLPH-TQAAPDDPLPLISDQTSPLSWFGLBEDDAVAEL-GIDFQRFLLTN 70
 66 GTIYIARDHIYVDIDTSH-T-EIYCSKLTWKSRAVDVTCMKHDECHNFIVL 124
 71 RTLLVAARDHVSPDLOAEDEEGELVFNKYLWRSQ--DVENCAVRKLTDECTNYIRVL 128
 125 LKKNDALFYCGTNAFNPSCRNVMDTLEPFGEFSGMACPYDAKANAVALPADGKLYS 184
 129 VPNDSCILLACGINSFSPVCRSYGITSLOQEGEELSGQACPFATQSNVAIFABGSLYS 188
 185 ATVTDLAIDAVIYRSLGSEPTLTFTVGHDSKMLKEPFVQAVDYGDIYFFPFIANEYN 244
 189 ATAADFQASDAVAVYRSIGPQPLRSAXYDSKWLREPHFVALBHGHDVFEFFREVSEDA 248
 245 TMGVVPPRYAVQVCKNMGSGQVLEKQWTSFLKARLNCSPGDSHFYFNILQAVTVIR 304
 249 RLGVQFSRVARVCKRMGSPRALDRHTSFLKRLNCSYFGDSTYFVLOALTGPNV 308
 305 INGRDVLAATFSTYNSIPGSAVCAVMDLADIASVFTGFKQKSPDSTWTPVDERVCKP 364
 309 LHGSALFGVTTQNTGSPSAVCAVFLDEIERGFEKQKQRLDGMATVSDRVPSP 368
 365 RPGCCAGSSSLERYATSNPEPDDTLNFKTHPLMDDAVPSIFNRPWFLRMVRLIKIA 424
 369 RFGSCAGVGAALPSSSRDLDDVLTFLKAPHLDPAPVPATHOP-LTLTTSRALLTQVA 427
 425 VDTAAGPQNTVFLGSEKGIILKFLARIGNSGFLNLSLEMSVYNSSEKSGDYVED 484
 428 VDMAGAPHSNTVFLGSDGTFLVKLTPGSRSG-PEPILBELIDAYSPPKCGKRAQ 486
 485 --KRIIMQDLRASSSLVAFSTCVIKVPLGRCEHCKKXTCIASRDPYCGMIKEGAC 542
 487 TARRIIGELDTGHRFLFAFSGCIYVPLRSCARHACQSRCLASQDPYCGMSSSRG-C 545
 543 SHLSPNSHLTFEODIERGNTDGL-GDCHNSFVALINGHSSSLPSTTSDTAQEGYSR 600
 546 VDIRSGSGTDVQ--AGNOSMEHGDQDQ--ATGSOQSGRGSAY-- 586
 601 GGMIDMKLLDSPOSTDPLGAVSSHNHODKGVIRESLYKGDHCLVPTLLAIAVILAFV 660
 587 -----GVRDLPPLASASRSVPILPLIASAALFA 615
 661 MGAVFSGITVYCVCDHRRKDVAVVQKKEKELTHSRG-----SMSSVTKLSGL 708
 616 LGASVSGILVSCAC--RR-----AHRRRKDIETPGPRLPLSLRSLARLHG- 659
 709 FGDTSQDKPP--EAILPLPMNGKLAFTGNTAKMLIKADQHLHLLTALPTPESTPTIQ 766
 660 -GGPPPPPSKDGDAVQTPOLYTTFLPPEGVPP-----ELACLPPESTPELAV 709
 767 KKKPSRSGSEMERONLINAACKMDPMPGSPVIFTDLPLRASPSH-----IPSVVLPIT 821
 710 KILRAAGD-PWEMNQRRNA-----KEGPRSGGHAAGAPARVILVRP- 752
 822 QCGYGEHYVDQPKMSEVAQMALEDQATLEYKTIKEHL-----SSKSPHGNVLVENDLS- 876
 753 -----PPGCGPG-----AVEVTLLELLRYLHGQPPRRGAGAPPALTSR 793
 877 -LPPKVPQREASLIGPPGASLSOTGLSKLEMHSSSYGVADYKRSYTNLSLTHSQATLX 935

DB 794 ALPPE--PAPALLGSPRPHRCASPLRLDY-----PPRCGASAPA----- 833
 QY 936 RNNVSSNSGHL-----SRNOSFGRGDNPPAPQRVDISIQHSSQPGCAVTVSRGS----- 988
 DB 834 -PPALGAPAPRLGVGGGRRLPFGHRAAPPALITRV-----PFGGSPRYSGGPGKHL 883
 QY 989 --LMAINSILTRSGILKRTPLSKPDVPEKPSF-APLSTSMKEN 1026
 DB 884 LVIGRPEGYRGALRKRVDEKFPQLSLKPLVGPBSSQAVPN 924

RESULT 4

US-09-254-594-3
 ; Sequence 3; Application US/09254594
 ; Patent No. 6566094
 ; GENERAL INFORMATION:
 ; APPLICANT: KIMURA, Toru
 ; APPLICANT: KIKUCHI, Kaoru
 ; TITLE OF INVENTION: NOVEL SEMAPHORIN GENE: SEMAPHORIN Y
 ; FILE REFERENCE: 0020-4527P
 ; CURRENT APPLICATION NUMBER: US/09/254,594
 ; CURRENT FILING DATE: 1999-05-11
 ; NUMBER OF SEQ ID NOS: 13
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 3
 ; LENGTH: 929
 ; TYPE: PRT
 ; ORGANISM: Rattus norvegicus
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: ()..()
 ; OTHER INFORMATION: Tissue Type: Brain
 ; LOCATION: ()..()
 ; OTHER INFORMATION: Identification Method: P for resulting peptide
 US-09-254-594-3

Query Match 26.2%; Score 1428.5; DB 4; Length 929;
 Best Local Similarity 34.08; Pred. No. 1.1e-126;
 Matches 360; Conservative 147; Mismatches 355; Indels 197; Gaps 32;

6 LLLYFTLLHPAGAFPEDESEPISSHGNYTKQYVFNVGHKRGKNTQRRHLDIQMIMIN 65
 11 LLLLSLPHQAQTAFFQDDPLPLISDQTSPLSWFGLBEDDAVAEL-GIDFQRFLLTN 69
 66 GTIYIARDHIYVDIDTSH-T-EIYCSKLTWKSRAVDVTCMKHDECHNFIVL 124
 70 RTLLVAARDHVSPDLOAEDEEGELVFNKYLWRSQ--DVENCAVRKLTDECTNYIRVL 127
 125 LKKNDALFYCGTNAFNPSCRNVMDTLEPFGEFSGMACPYDAKANAVALPADGKLYS 184
 128 VPNDSCILLACGINSFSPVCRSYGITSLOQEGEELSGQACPFATQSNVAIFABGSLYS 187
 185 ATVTDLAIDAVIYRSLGSEPTLTFTVGHDSKMLKEPFVQAVDYGDIYFFPFIANEYN 244
 188 ATAADFQASDAVAVYRSIGPQPLRSAXYDSKWLREPHFVALBHGHDVFEFFREVSEDA 248
 245 TMGVVPPRYAVQVCKNMGSGQVLEKQWTSFLKARLNCSPGDSHFYFNILQAVTVIR 304
 249 RLGVQFSRVARVCKRMGSPRALDRHTSFLKRLNCSYFGDSTYFVLOALTGPNV 308
 305 INGRDVLAATFSTYNSIPGSAVCAVMDLADIASVFTGFKQKSPDSTWTPVDERVCKP 364
 309 LHGSALFGVTTQNTGSPSAVCAVFLDEIERGFEKQKQRLDGMATVSDRVPSP 368
 365 RPGCCAGSSSLERYATSNPEPDDTLNFKTHPLMDDAVPSIFNRPWFLRMVRLIKIA 424
 369 RFGSCAGVGAALPSSSRDLDDVLTFLKAPHLDPAPVPATHOP-LTLTTSRALLTQVA 427
 425 VDTAAGPQNTVFLGSEKGIILKFLARIGNSGFLNLSLEMSVYNSSEKSGDYVED 484
 428 VDMAGAPHSNTVFLGSDGTFLVKLTPGSRSG-PEPILBELIDAYSPPKCGKRAQ 486
 485 --KRIIMQDLRASSSLVAFSTCVIKVPLGRCEHCKKXTCIASRDPYCGMIKEGAC 542
 487 TARRIIGELDTGHRFLFAFSGCIYVPLRSCARHACQSRCLASQDPYCGMSSSRG-C 545
 543 SHLSPNSHLTFEODIERGNTDGL-GDCHNSFVALINGHSSSLPSTTSDTAQEGYSR 600
 546 VDIRSGSGTDVQ--AGNOSMEHGDQDQ--ATGSOQSGRGSAY-- 586
 601 GGMIDMKLLDSPOSTDPLGAVSSHNHODKGVIRESLYKGDHCLVPTLLAIAVILAFV 660
 587 -----GVRDLPPLASASRSVPILPLIASAALFA 615
 661 MGAVFSGITVYCVCDHRRKDVAVVQKKEKELTHSRG-----SMSSVTKLSGL 708
 616 LGASVSGILVSCAC--RR-----AHRRRKDIETPGPRLPLSLRSLARLHG- 659
 709 FGDTSQDKPP--EAILPLPMNGKLAFTGNTAKMLIKADQHLHLLTALPTPESTPTIQ 766
 660 -GGPPPPPSKDGDAVQTPOLYTTFLPPEGVPP-----ELACLPPESTPELAV 709
 767 KKKPSRSGSEMERONLINAACKMDPMPGSPVIFTDLPLRASPSH-----IPSVVLPIT 821
 710 KILRAAGD-PWEMNQRRNA-----KEGPRSGGHAAGAPARVILVRP- 752
 822 QCGYGEHYVDQPKMSEVAQMALEDQATLEYKTIKEHL-----SSKSPHGNVLVENDLS- 876
 753 -----PPGCGPG-----AVEVTLLELLRYLHGQPPRRGAGAPPALTSR 793
 877 -LPPKVPQREASLIGPPGASLSOTGLSKLEMHSSSYGVADYKRSYTNLSLTHSQATLX 935

```

REFERENCE/DOCKET NUMBER: B94-002-1
TELEPHONE: (415)343-4341
TELEFAX: (415) 343-4342
TELEX:
INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 730 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-121-713D-58

Query Match 17.8%; Score 969.5; DB 1; Length 730;
Best Local Similarity 34.4%; Pred. No. 4,56-83;
Matches 244; Conservative 124; Mismatches 224; Indels 117; Gaps 29

      1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999
```

```

APPLICANT: Matthes, David
APPLICANT: Bentley, David R.
APPLICANT: O'Connor, Timothy
TITLE OF INVENTION: The Semaphorin Gene Family
NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESS:
ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 Bush Street, Suite 3200
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/835,268
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/06/121,713
FILING DATE: 13-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B94-002-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)343-4341
TELEFAX: (415) 343-4342
TELEX:
INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 730 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-835-268-58

Query Match      17.8%; Score 969.5; DB 1; Length 730;
Best Local Similarity 34.4%; Pred. No. 4.5e-83;
Matches 244; Conservative 124; Mismatches 224; Indels 117; Gaps 29;

QY 8 LFTLLFAGAGFPEDESPISISHGNYTKQYVFVGHKPGKNTQR-----HRLDIQMI 61
DB 11 LLMVALH-AAAMNDVSP-----KMYVQF-----GERVQRFIGNESHKHFKL 54
QY 62 MINGGTLYIARHIIYTVDD--TSHTEIYCSKLTWKSQADVDTCRMKSKHDEBN 119
DB 55 EKDNHSLVAGARNTVYNISLRDLTEFTQ-----RIEWHSSGAHRELCTYLGKSEDDCN 109
QY 120 FIKYLLKNDNALFVCGTNAFNPSCRYNKMDTLEPFGD-----EFGMARCPYAKHANY 174
DB 110 YIRVIAKIDDDRVLICGTNAKPLCRHIALD-----GIYVVEKEVEGGLCPFPBDHST 165
QY 175 ALFADGKLYSATVTDPLAIDAVIYKSLGSEPTLRTVKHDSKMLKEPYVQAVDGYIYF 234
DB 166 AIYSGQIYSATVADFSGTDLIYRG-----PLRTERSDLKQINAPFNVTMEYVDFEYF 220
QY 235 FFRREIAYENMKGVFPRVAVQVCKNDMGSSORVLEKQMTSTLKLRLNCVPGDSHEFEN 294
DB 221 FFRREIAYENMKGVFPRVAVQVCKNDMGSSORVLEKQMTSTLKLRLNCVPGDSHEFEN 279
QY 295 ILQAVTVIRLN-GADV---VLAFTSTPYNISPGSAVADMLDIASVFTGRFKEQKSPD 350
DB 280 EIQSTSDIIEGNYGQVEKLIYGVFTTPNLSIGSAVCAFSKSLLESPDGFKEQETMN 339
QY 351 STMTVPDEPRVYKPPGCCAGSSSLERYATSNFPDDTLNFKTLPDLNDEAVESIFNRPW 410
DB 340 SNMLAVPSLKVDEPRPGQCVND-----SRLPDVSVNFVSKSHTLNDEAVFAFTBPI 391
QY 411 FLRTWVRIRLTKIAND---TAAGPYQNHVTVFQSEKIIILKFLARIGNSGFLN----- 461

```

```

DB 392 LRISLQYRFTIANDQVTRPDG--KAYDVLFIGDDDKVIAL-----NSASFDSSTV 445
QY 462 DELFLEMSVYNSEKSYDGEDKRIWQOLDRASSLVANSTCYIKVPLGC--ERHG 519
DB 446 DSVVIELQVLP-----PGVPVKLYVRMGDDSKLVVSDDELIAIKIHRGSDKIT 499
QY 520 KCKKTIASRPDYCGMIKRGACSHL-SPN-----SRLTFQDIERGNTDGLGCHNSFVAL 575
DB 500 NCRE-CVSLQDPYCAWMDVNLKCTAVGSPDWAGKRRFTQISLGHKACGRRPQTEIV- 557
QY 576 NGHSSSLPSTTSDSTA-----QEGYERGMWMDKHLDSPDSTPLGAVSSHNQ 628
DB 558 ---ASPVPTRQPTTKSGDPVSHQAEPPE---IDNEIVIGVDSNVNPTLAEINHA 610
QY 629 DKGVIRESYLKGHDQVLPV---TL-LAI-----AVIAPYMGVFS 666
DB 611 GSK-----LPSSQKELPYTAETLTIAIVISCLGALVVGFIISGLFS 652

```

```

RESULT 7
US-09-060-692-58
Sequence 58, Application US/09060692
Patent No. 5935865
GENERAL INFORMATION:
APPLICANT: Goodman, Corey S.
APPLICANT: Kolodkin, Alex L.
APPLICANT: Matthes, David
APPLICANT: Bentley, David R.
APPLICANT: O'Connor, Timothy
TITLE OF INVENTION: The Semaphorin Gene Family
NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESS:
ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 Bush Street, Suite 3200
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/060,692
FILING DATE:
CLASSIFICATION: 514
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/08/121,713
FILING DATE: 13-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B94-002-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)343-4341
TELEFAX: (415) 343-4342
TELEX:
INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 730 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-060-692-58

Query Match      17.8%; Score 969.5; DB 2; Length 730;
Best Local Similarity 34.4%; Pred. No. 4.5e-83;
Matches 244; Conservative 124; Mismatches 224; Indels 117; Gaps 29;

QY 8 LFTLLFAGAGFPEDESPISISHGNYTKQYVFVGHKPGKNTQR-----HRLDIQMI 61

```

```

Db      11 LLMVALH-AAAMVNDVSP-----KMYVQF-----GEEVQRFIGNESHKHDFKLL 54
QY      62 MIMNGTLYIARDDIYTVDID--TSHTEEIYCSKKLTWKSROADVDTRMGKHKDECHN 119
Db      55 EKDHSLSLVGARNIYINISLRLTFTEQ-----RIEHSAGAHRELCTYLGKSSDDCN 109
QY      120 FIKVLLKKNDLDFVCGTNAFNSCRNYKMDTLEPFGD-----EFSGARCPYDAKHANV 174
Db      110 YIRVLAKIDDDRVLLCGTNAKPLCRHYALKD-----GDYVEKEVEYEGGLCPFDPHNST 165
QY      175 ALFADGKLYSATVTPFLAIDAVIYSLGESPLRTVKHDSKMLKEPYQAVADYDGYIYF 234
Db      166 AYESGOLYSATVADFSGTDLPIYRG-----PLKTERSDLKQLNAPNVTNMEYNDFTLF 220
QY      235 FPREIAVENYTMGKVVPRPAQVCKNDMGSGQVLEKQWTSFLKARLNCVSGDSHFYFN 294
Db      221 FPREIAVEYINGCKAIYSKVARVCHDKGPHQGDGR--WTSFLKRLNCVSGDYPFFYN 279
QY      295 ILQAVTVIRIN-GRDV--VLATSTPYNSIPGSAVCAVMDLIDASVFTGRFEKQKSPD 350
Db      280 EIQSTSDIIEGNYGGVEKLIYGVFTTPVNSIGSAVCAFSKMSILSEFDGPFKEQETWN 339
QY      351 STWTPVPDEVRVPRPGCCAGSSSLERVATSNFPPDITNFIKTHPLMDEAVPSIFNRPW 410
Db      340 SNMLAVPSLKVPEPRPGQCVND-----SRTLPPVSVNPFVKSHTLMDAIVAPAFTRPI 391
QY      411 FLRTWVRVRLTKIADV-----TAAGPYQNTVFLGSEKGIILKFLARIGNSGFLN----- 461
Db      392 LRISLQYFTKIAVQVQVTRPDG--KAYDVLFIGTDDGKVIKAL-----NSASFSSDTV 445
QY      462 DSLFLEEMSVYNSEKSYDGEDKRIWGQOLDRASSLYVANSTCYIKVPLGRC--ERRG 519
Db      446 DSVVIELQVLP-----PGVPYKMLYVVRMDGDGSKLVVSDDELLAIKLHRCGSDKIT 499
QY      520 KCKKTCTIAGSDPYCGWIKEGGACSHL--SPN-----SRLTFEODIERGNTDGLGDCHNSFVAL 575
Db      500 NCRE--CVSLQDPYCAMDNVELKCTAVGSPWSAGKRFFIONISLGEHKAAGGRPQTEIV- 557
QY      576 NGHSSSLPSTTSDDSTA-----OEGVESRGMGLDWKHLDPSTPLGAVSSHNO 628
Db      558 ----ASPVPQPTTKSSGDPVSHIQAEPEPE--IDNEIVIGVDSNVITPNTLAINHA 610
QY      629 DKKGIVRESYLKGDQLVPPV---TL-LAI-----AVLIAFVMAVFS 666
Db      611 GSK-----LPSSQEKLPITYTAETLITAIIVTSLGLALVVGFSIGFLFS 652

RESULT 8
US-08-833-391-58
; Sequence 58, Application US/08833391
; Patent No. 6013781
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matches, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/833,391

```

```

; FILING DATE: 530
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/121,713
; FILING DATE: 13-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
;
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 730 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
;
; MOLECULE TYPE: protein
;
; US-08-833-391-58
;
; Query Match 17.8%; Score 969.5; DB 3; Length 730;
; Best Local Similarity 34.4%; Pred. No. 4.5e-83;
; Matches 244; Conservative 124; Mismatches 224; Indels 117; Gaps 29;
;
;
; 8 LYFTLHPAGAGFPEDSEPISSHGNTYKQVVFVGHKPGANTQR-----HRLDIOMI 61
;
; Db      11 LLMVALH-AAAMVNDVSP-----KMYVQF-----GEEVQRFIGNESHKHDFKLL 54
QY      62 MIMNGTLYIARDDIYTVDID--TSHTEEIYCSKKLTWKSROADVDTRMGKHKDECHN 119
Db      55 EKDHSLSLVGARNIYINISLRLTFTEQ-----RIEHSAGAHRELCTYLGKSSDDCN 109
QY      120 FIKVLLKKNDLDFVCGTNAFNSCRNYKMDTLEPFGD-----EFSGARCPYDAKHANV 174
Db      110 YIRVLAKIDDDRVLLCGTNAKPLCRHYALKD-----GDYVEKEVEYEGGLCPFDPHNST 165
QY      175 ALFADGKLYSATVTPFLAIDAVIYSLGESPLRTVKHDSKMLKEPYQAVADYDGYIYF 234
Db      166 AYESGOLYSATVADFSGTDLPIYRG-----PLKTERSDLKQLNAPNVTNMEYNDFTLF 220
QY      235 FPREIAVENYTMGKVVPRPAQVCKNDMGSGQVLEKQWTSFLKARLNCVSGDSHFYFN 294
Db      221 FPREIAVEYINGCKAIYSKVARVCHDKGPHQGDGR--WTSFLKRLNCVSGDYPFFYN 279
QY      295 ILQAVTVIRIN-GRDV--VLATSTPYNSIPGSAVCAVMDLIDASVFTGRFEKQKSPD 350
Db      280 EIQSTSDIIEGNYGGVEKLIYGVFTTPVNSIGSAVCAFSKMSILSEFDGPFKEQETWN 339
QY      351 STWTPVPDEVRVPRPGCCAGSSSLERVATSNFPPDITNFIKTHPLMDEAVPSIFNRPW 410
Db      340 SNMLAVPSLKVPEPRPGQCVND-----SRTLPPVSVNPFVKSHTLMDAIVAPAFTRPI 391
QY      411 FLRTWVRVRLTKIADV-----TAAGPYQNTVFLGSEKGIILKFLARIGNSGFLN----- 461
Db      392 LRISLQYFTKIAVQVQVTRPDG--KAYDVLFIGTDDGKVIKAL-----NSASFSSDTV 445
QY      462 DSLFLEEMSVYNSEKSYDGEDKRIWGQOLDRASSLYVANSTCYIKVPLGRC--ERRG 519
Db      446 DSVVIELQVLP-----PGVPYKMLYVVRMDGDGSKLVVSDDELLAIKLHRCGSDKIT 499
QY      520 KCKKTCTIAGSDPYCGWIKEGGACSHL--SPN-----SRLTFEODIERGNTDGLGDCHNSFVAL 575
Db      500 NCRE--CVSLQDPYCAMDNVELKCTAVGSPWSAGKRFFIONISLGEHKAAGGRPQTEIV- 557
QY      576 NGHSSSLPSTTSDDSTA-----OEGVESRGMGLDWKHLDPSTPLGAVSSHNO 628
Db      558 ----ASPVPQPTTKSSGDPVSHIQAEPEPE--IDNEIVIGVDSNVITPNTLAINHA 610
QY      629 DKKGIVRESYLKGDQLVPPV---TL-LAI-----AVLIAFVMAVFS 666
Db      611 GSK-----LPSSQEKLPITYTAETLITAIIVTSLGLALVVGFSIGFLFS 652

```

RESULT 9
US-09-060-610-58
Sequence 58, Application US/09060610
Patent No. 6344544
GENERAL INFORMATION:
APPLICANT: Goodman, Corey S.
APPLICANT: Kolodkin, Alex L.
APPLICANT: Matthes, David R.
APPLICANT: Bentley, David R.
APPLICANT: O'Connor, Timothy
TITLE OF INVENTION: The Semaphorin Gene Family
NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY IAW GROUP
STREET: 268 Bush Street, Suite 3200
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/060,610
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/835,268
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B94-002-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
TELEX:
INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 730 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-060-610-58

Query Match 17.8%; Score 969.5; DB 4; Length 730;
Best Local Similarity 34.4%; Pred. No. 4.5e-83;
Matches 244; Conservative 124; Mismatches 224; Indels 117; Gaps 29;

QY 8 LYFTLLHRAAGFPDSEPTISHSNNTKQYVFGHKGKNTTOR-----HRLDIOMI 61
DB 11 LHMVALH-AAAMWADVSP-----KMYVQF-----GERRVQRELGESHKDFEKL 54
QY 62 MIMNGTLYIARBDHLYTVDID--TSHTEIYCSKLTWKSROADVDTCKMGKHNDECHN 119
DB 55 EKDNHSLVAGARNIYINISLRDLTFTEO-----RIEMHSSAHNELCYLKGKSDDCQN 109
QY 120 FIKVLLKKNDDALFVCGTNAFNPSCRNKMDTLEPFGD-----EFGMARCPYDAKMANV 174
DB 110 YIRVLAKTIDDDRVLLCGNNAKPLCRHYALKD-----GDYVVEKEVEGKGLCFDPDHNST 165
QY 175 ALFADGKLYSATVTFPLAIDANITYSLGESPTLFTVKHDSKYLKBPYFQAVDGYCYTYF 234
DB 166 AYSGGQYLSATVAOFSGTDPITYRG-----PRTERSDLKOLAPNPNVMTMEINDIFF 220
QY 225 FPRELAVENYMGKVFEPVAVQCKNDMGSGORVLEKQWTSFLKARLNSVGDGSHFYFN 294
DB 221 FPRELAVENYMGKVFEPVAVQCKNDMGSGORVLEKQWTSFLKARLNSVGDGSHFYFN 279
QY 295 ILQAVTVDVIRIN-GRDV--VIALTFSTPYNSIPGSAVCAVDMDLIASVTFGTGFKQKSPD 350

DB 280 EIQSTSDIIEGNYGQVEKLIYVFTTPVNSIGSAAVCAFPKSKSLIESDGPKEGETVN 339
QY 351 STWTPVDEPDERVPPKPPGCCAGSSSLERYATSNFEDDTLINF1KTHPLMDEAVPSINRPW 410
DB 340 SNMLAVPSLAKVEPFRPGQVND-----SRTLDPVSNVFPVSRMLDEAVPAFTRPI 391
QY 411 FLRTMVRVRLTKIAYD---TAAGPYONHTVVFLSEKGLILKFLARIGNSGFLN----- 461
DB 392 LIRISLQYRFTKIAVDQOVRTPDG--KAYDVLFIGDDCKVIAL---NSASFDSQDTV 445
QY 462 DSLFLEEMVSNSEKSYGVDEKRIIMGQDLRASSSLVAFSTCYIKVPLGRG--ERHG 519
DB 446 DSVVIEELQVLP-----PQVYKXLYVVRMGDDSKLVVSGDELLAIKLNHCGSDKIT 439
QY 520 KCKTICIASRDPYCGWIKGACSHL-SPN---SRLTFQDIERGNTDGIQDCNSFVAL 575
DB 500 NCRF-CVSLQDPYCANVDNELKCTAVGSDWSAGKRFQNLISGHHKACGRPQTEIV- 557
QY 576 NHHSSLLPSTTSDSTA-----QEGYESRGMDMKHLDSPSTDPDLGAVSSHNNQ 628
DB 558 ---ASPVPTQPTKSSGDPVHSIQAEPEPE---IDNEIVIGVDVSNVPIPTLAEINHA 610
QY 629 DKXGVIRBSYKGDQLVPY---TL-LAI-----AVILAFVMAVFS 666
DB 611 GSK-----LPSSQEKPLPYTAETLITLIALVTSCLGALVVGFTSGFLFS 652

RESULT 10
PCT-US94-10151A-58
Sequence 58, Application PC/TUS9410151A
GENERAL INFORMATION:
APPLICANT: The Regents of the University of California
TITLE OF INVENTION: The Semaphorin Gene Family
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR HOBBACH TEST ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/10151A
FILING DATE: 13-SEP-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: FP-58750-PC/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299 FHT UR
INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 730 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US94-10151A-58

Query Match 17.8%; Score 969.5; DB 5; Length 730;
Best Local Similarity 34.4%; Pred. No. 4.5e-83;
Matches 244; Conservative 124; Mismatches 224; Indels 117; Gaps 29;

QY 8 LYFTLLHRAAGFPDSEPTISHSNNTKQYVFGHKGKNTTOR-----HRLDIOMI 61

```

Db      11 LLMVALH-AAAMVNDVSP-----KMYVQF-----GEERVOPLGNSHDKHEKLL 54
QY      62 MIMNGLYIARADHIYTDID--TSHTTEIYCSKLLTWKSPQADVDTCKMKHDECHN 119
Db      55 EKDHNSLVGANKYIVINISRLDITETEQ-----RIEMHSSAHNELCTLKKSSEDCCN 109
QY      120 FIKVLKKNDDALFVCGTNAFNPSCNRYMDTLEPPGD-----EFGMARCPYDAKHANV 174
Db      110 YIRVLAKIDDDVLLIGTNAYKPLGRHYALKD-----GDYVEKEVEGRGLCPFPDDHNS 165
QY      175 ALFADGKLYSATVTDPLAIDAVIYRSLGSSPLIRYKHDSKYLKEFYFOADYGDYIYF 234
Db      166 AIYSEGLYSATVADPSGTDPLIYRG-----PLKTERSDLKOLNMFNTVEYNDYDFLF 220
QY      235 FPREIYEVNTMGKVPFPPVAQVCKNDMGSGORVLEKOWTSPFLKARLNCVSGDSHFYFN 294
Db      221 FPREIYEVNTMGKVPFPPVAQVCKNDMGSGORVLEKOWTSPFLKARLNCVSGDSHFYFN 279
QY      295 ILQAVTAVIRIN-GRDV---VLATFSTPYNSIPGSANCAVMDLILASVFTGPRKQKSPD 350
Db      280 ELQSTSDIIEGNVGGQVEKLIYGVFTTPYNSIGSAAVCAFSMKSLIESFDGPFKEQETWN 339
QY      351 STWTPYDPERVPEKPRGCCAGSSSLERVATSNFPDPTLNFKTHPLMEAVPSIENRW 410
Db      340 SNMLAVPSLKVPEPRGQCVND-----SRTLPEVSNVFNKSHLMDPAVAFTRFI 391
QY      411 FLATWRYRLTKLAVD-----TAAGPYQNHVTVFLGSEKGIILKFLARIGNSGFLN----- 461
Db      392 LIRISIQYRFTKLAIVQOVRTPDG--KAYDVLEFIGDDGKVIKAL-----NSASFDSQDVT 445
QY      462 DSLFLEMSVYNSEKSYDGEDKRIHQMDRASSSLVYAFSTCYIKPPLGRC--ERBG 519
Db      446 DSVVIEELQVLP-----PGVPKNLYVVMDDGDKLIVVSDDEILAIKLHRCGSDKIT 499
QY      520 KCKKTCLASDPPYCGMIKEGAGASHL-SPN---SRLTFEODIERGNTDGLGCHNSFVL 575
Db      500 NCRE-CVSLDDPYCAMDNVLELCTAGSSPWSAGKRREFIONISLGHHKACGGRPOTEIV 557
QY      576 NGHSSSLPSTTSDSTA-----OEGYBSRGGMIDMKILDSPSTDPGLAVSSHNQ 628
Db      558 ---ASVPTQPTTKSSGDPVSHIHQAFEPPE---IDNELVIGVDDSDNV.PNTLAEINHA 610
QY      629 DKKGVIRESYLKGHDLVPTV---TL-LAI-----AVIIAFVWGAVER 666
Db      611 GSK-----LPSSQEKLPYTAETTLITAIYVTSGLALVVGFLSGFLFS 652

```

```

RESULT 11
US-08-121-713D-60
; Sequence 60, Application US/08121713D
; Patent No. 5639856
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/121,713D

```

```

; FILING DATE: 13-SEP-1993
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)343-4341
; TELEFAX: (415) 343-4342
; TELEX:
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 650 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-121-713D-60

```

```

Query Match      16.1% Score 880; DB 1; Length 650;
Best Local Similarity 35.8% Pred No. 1.3e-74;
Matches 207; Conservative 95; Mismatches 194; Indels 82; Gaps 19;

QY      114 KDCHENFLKVLKKNDDALFVCGTNAFNPSCNRYMD-----TLPPGDEFGMARCPYNA 169
Db      1 EDDCQNYIRIMVPSPPGRFLFVCGTNSPPMCNTYIISDSNYTLIA---TKNGQAVCPYDP 57
QY      170 KXANVALFADGKLYSATVTDPLAIDAVIYRSLGSEPTLRTHGDSKYLKEFYFOADVYG 229
Db      58 RHNSTSVIADNDELVSGTAVDFSGSDPIIYRE-----PLQTEQYDLSLAAFPVSSFTQG 112
QY      230 DYIYFPPREIAYEVNTMGKVPFPPVAQVCKNDMGSGORVLEKOWTSPFLKARLNCVSGDS 289
Db      113 DSVVIEFRTAVLEFINGGAIYSRVARVCKMKGGPR--FRNRWTSFLSRINCSPIDGY 171
QY      290 HEYFNILQAVTVIR---INGRDVVLATFSTPYNSIPGSANCAVMDLILASVFTGPRERE 345
Db      172 PFYFMEIQASNLVEQYGSMSKLIYGVFTTPYNSIPGSANCAVMDLILASVFTGPRERE 231
QY      346 QKSPDSTWTPVDEDERPKPRGCCAGSSSLERVATSNFPDPTLNFKTHPLMEAVPSI 405
Db      232 QGINSNMLPVNNAKVDFPRPSC-----HNSRALPDPLNFKTHSLMDENVAPF 283
QY      406 FNRPFILRTMVRVRLTKLAVD-----TAAGPYQNHVTVFLGSEKGIILKFL-ARIGNSGFL 460
Db      284 FSPILVIRSTIYRFQIAVDAQIKTPG--KTYDVIFVGTGDKLIKSYNAESADSADK 341
QY      461 NDSLFEEMSVYNSEKSYDGEDKRIH-QMOLDRASSSLY-----VAFSTVIRKPLG 513
Db      342 VTSVIEIEIDVLTKS-----EPINLEIVRTMOYDQPKDGSYDDGKLIIVTDSQVVAIQH 397
QY      514 RC--ERHGCKKCTCIASRDPYCGMIKEGAGC-SHLSPN---SRLTFEODIERGNTDGLGDC 568
Db      398 RCHNDKITSCS-CVALDDPYCAMDKIAGCRSHGAPRWLEENYFYQNVATGQ----- 449
QY      569 HNSFVALNGHSSSLPSTTSDSTAQEGYBSRGGMIDMKILDSPSTDPGLAVSSHNQ 628
Db      450 -----HAACPSGKINSKDANAGEQKFRNDM-----DLIDS-----RRQ 483
QY      629 DKKGVIRESYLKGHDLVPTVTLAIVIIAFVWGAVER 666
Db      484 SKDQEIINDIKNFEDIINAGYVETIIVMAVLGSLFS 521

```

```

RESULT 12
US-08-835-268-60
; Sequence 60, Application US/08835268
; Patent No. 5807826
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy

```



```

Db 172 PFYFNEIQSASNLVEGQYGSMSKLIYGVFNTSPNSIPGSAVCAFALQDIADTFEGQFKE 231
Qy 346 QKSPDSTWTPVDERVPRKPPGCCAGSSSLERVATSNFPDDTLNFTKTHPLMDEAVPSI 405
Db 232 QTGINSNMLPVNNAKVPDRPGSC-----HNSRALPPTLNFTKTHSLMDENVPF 283
Qy 406 ENRPFELTMRVYRRTKTAVD-----TAAGPYONHTVFLGSEKGIILKFL-ARIGSGFL 460
Db 284 FSQPLVTRTSTYRFTQIAVDAQITPGG--KTYVIVIVGIDHGKLIISVNAESADSDX 341
Qy 461 NDSLFLFEMSVYNSKSCYDGVEDKRIIM-GMQLDRASSSLY-----VAFSTCVIKVPLG 513
Db 342 VTSVIVIEIDVLTGS-----EPINLEIVRTWQYDQPKGSDYDDGKLIIVTDSQVAIAQLH 397
Qy 514 RC-ERHGKCKKTCCTASDPYCGWIKEGAC-SHLSPN--SLTFEODIERGNTDGLGDC 568
Db 398 RCHNDKITSQSC-CVALDPPYCAMDKIAGKCRSHGAPRWLENNYVQNAVATQ-----449
Qy 569 HNSFVALNGHSSSLPSTTSDSTAOEGYESRGMIDWGHLLDSPDSTPLGAVSSHNO 628
Db 450 -----HAACPSGKINSKDNANGEGQGFNDX-----DLDS-----RRQ 483
Qy 629 DKGVIRSESYLKGHDLVPTLLAIAVILAFVWGAVFS 666
Db 484 SKDQELIIDNIDNFDIINAQYTVETLVAVLAGSIFS 521

RESULT 14
US-08-833-391-60
; Sequence 60, Application US/08833391
; Patent No. 6013781
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David R.
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/833,391
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/121,713
; FILING DATE: 13-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)343-4341
; TELEFAX: (415) 343-4342
; TELEX:
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 650 amino acids
; TYPR: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

```

```

US-08-833-391-60
Query Match 16.1%; Score 880; DB 3; Length 650;
Best Local Similarity 35.8%; Pred. No. 1,3e-74;
Matches 207; Conservative 95; Mismatches 194; Indels 82; Gaps 19;

Qy 114 KDEHNFIIXVLKKNDDALFVCGNMFNPGCRNYKMD-----TLFPDEFSGMARCFYDA 169
Db 1 EDDCQYIRLWVDPSPGRFLVCGINSRPMCNITYIIDSNTTLEA-----TKNGQAVCPDP 57
Qy 170 KHANVALFADGKLVSATVDTFLAIDAVIYRSLGSPFLRTVYKDSKWLKEPYFOAVDYG 229
Db 58 RHNSTVLADNELISGVTVADPSGSDPIIYRE-----PLQTEQYDLSLAINAFVSSFTQG 112
Qy 230 DYIYFFPREIAVEYNMGKVPFRVAYQYCKNDMGSGQVLEKQMTSPFKARLNCSPGDS 289
Db 113 DVAFFFEFRETAVEFINCGKALYSRVARCKWDKGGPHR-FENRMTSPKSLNCSIPDXY 171
Qy 290 HVENIILQAVTVDR-----INGRDVLAFTSTPYNSIPGSAVCAVMDLIDIASVFTGRPE 345
Db 172 PFYFNEIQSASNLVEGQYGSMSKLIYGVFNTSPNSIPGSAVCAFALQDIADTFEGQFKE 231
Qy 346 QKSPDSTWTPVDERVPRKPPGCCAGSSSLERVATSNFPDDTLNFTKTHPLMDEAVPSI 405
Db 232 QTGINSNMLPVNNAKVPDRPGSC-----HNSRALPPTLNFTKTHSLMDENVPF 283
Qy 406 ENRPFELTMRVYRRTKTAVD-----TAAGPYONHTVFLGSEKGIILKFL-ARIGSGFL 460
Db 284 FSQPLVTRTSTYRFTQIAVDAQITPGG--KTYVIVIVGIDHGKLIISVNAESADSDX 341
Qy 461 NDSLFLFEMSVYNSKSCYDGVEDKRIIM-GMQLDRASSSLY-----VAFSTCVIKVPLG 513
Db 342 VTSVIVIEIDVLTGS-----EPINLEIVRTWQYDQPKGSDYDDGKLIIVTDSQVAIAQLH 397
Qy 514 RC-ERHGKCKKTCCTASDPYCGWIKEGAC-SHLSPN--SLTFEODIERGNTDGLGDC 568
Db 398 RCHNDKITSQSC-CVALDPPYCAMDKIAGKCRSHGAPRWLENNYVQNAVATQ-----449
Qy 569 HNSFVALNGHSSSLPSTTSDSTAOEGYESRGMIDWGHLLDSPDSTPLGAVSSHNO 628
Db 450 -----HAACPSGKINSKDNANGEGQGFNDX-----DLDS-----RRQ 483
Qy 629 DKGVIRSESYLKGHDLVPTLLAIAVILAFVWGAVFS 666
Db 484 SKDQELIIDNIDNFDIINAQYTVETLVAVLAGSIFS 521

RESULT 15
US-09-060-610-60
; Sequence 60, Application US/09060610
; Patent No. 6344544
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,610

```

FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/835,268
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B94-002-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)343-4341
TELEFAX: (415) 343-4342
TELEX:
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 650 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-060-610-60

Query Match 16.1%; Score 880; DB 4; Length 650;
Best Local Similarity 35.8%; Pred. No. 1.3e-74;
Matches 207; Conservative 95; Mismatches 194; Indels 82; Gaps 19;

QY 114 KDECNFIKVLKKNDALFVCGTNAFNPSCRNKMD---TLPEFGDFSGMARCPYDA 169
DB 1 EDDCGNYIRIMVPSGRFLFVCGTNSFRPMCTYIIISDNYTLFA---TKNGQAVCPYDP 57

QY 170 KHAVALPADGKLYATYTDFAIDAVIYRSLGESPTLRIVYHDSKMLKEPFIQAVDYG 229
DB 58 RHNSSTVALADNELYSGTVADFGSDPIIYRE---PLQTEQYDSLNLAPFVSSFTQG 112

QY 230 DVIYFPEEELAEVNTMGVVFPRVAQVQNDMGSGQRYLEKQWTSFLKARLNCVPGDS 289
DB 113 DVYFPEEELAEVNTMGVVFPRVAQVQNDMGSGQRYLEKQWTSFLKARLNCVPGDS 289

QY 290 HEYFNILQAVMDVIR---INGRDVLAFTSTPYNISPGSAVCAYMDLIDIASVFTGRFE 345
DB 172 PFYFNEIQSASNLVGGQSGMSKLIYGVNTFNSIPGSAVCAPALODIADTFEGQFE 231

QY 346 QKSPDSTWTPVDPDERVPRPCCGAGSSSLERYATSNFPPDITNFIKTHPLMDEAVPSI 405
DB 232 QTGINSNMLPVNNAKVPPDRPGSC-----HNSRALDPPTLNFIKTHSLMDENVPAF 283

QY 406 ENRPWFLETRMAYRRLTKIADV---TAAGPYQNHVTVFLGSEKGIILKPL-ARIGNSGFL 460
DB 284 FSQPLIVRTSTIYRFTQIADVAQIKTPGG--KTYDVIYFVGTDHGKLIKSVNAESADSLDK 341

QY 461 NDSLPLEMSVYNSEKSYDVEDKRIK-GMQLDRASSLY-----VAFSTCVIKVPLG 513
DB 342 VRSVVIIEIDVLTKS---EPIRNLIEIVRTMQYDQKDSYDDGKLIIVTDSQVAIQLH 397

QY 514 RC-ERHGKCKTCTIASRPDYGMIKEGAGC-SHLSFN--SRUTEQDIERGNTDGLGDC 568
DB 398 RCHNDKITSCSE-CVALQDPYCAMDKIACKRSHGAPRWLEENYFYQNAVATGQ----- 449

QY 569 HNSFVALNGHSSSLDPTSDSTAQEGYESRGMIDMKHLDSPDSTDLGAVSSHNQ 628
DB 450 -----HAACPSGKINSKDNAGEQKGRNDM-----DLIDS-----RRQ 483

QY 629 DKKGVIRESYLKGDQIVPVTLLAIVLIAVFMGAVFS 666
DB 484 SKDQETIDNDKNEFDIINNAQYVETLVAVLAGSIFS 521

Search completed: October 23, 2003, 17:10:59
Job time: 23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: October 23, 2003, 17:09:07 ; Search time 26 Seconds
(without alignments)
3809.759 Million cell updates/sec

Title: US-09-856-681-2

Perfect score: 5450

Sequence: 1 MRSEALLLYTLHFAGAGF.....PPKSPAPLSTSMKNDPACT 1030

Scoring table: BLOSUM62

Searched: 283308 segs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:
2: pir2:
3: pir3:
4: pir4:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	971.5	17.8	730	2 JH0798	fasciclin IV precu
2	875	16.1	656	2 B49423	semaphorin I - fru
3	852.5	15.6	771	2 D49423	semaphorin III preu
4	842.5	15.5	711	2 A49423	semaphorin I preu
5	839.5	15.4	772	2 I48747	semaphorin D - mou
6	830	15.2	772	2 A49069	collapsin - chicke
7	826	15.2	1074	2 JCS928	semaphorin F preu
8	826	15.2	749	2 G01856	semaphorin V - hum
9	803	14.7	748	2 I48744	semaphorin A - mou
10	793.5	14.6	666	2 I58169	semaphorin III - m
11	789	14.5	712	2 I27165	hypothetical prote
12	779.5	14.3	724	2 C49423	semaphorin II preu
13	768	14.1	753	2 G02173	semaphorin III fam
14	737	13.5	751	2 I48748	semaphorin E - mou
15	692	12.7	834	2 S66498	M-sema F protein p
16	656	12.0	782	2 I48746	semaphorin C - mou
17	641.5	11.8	760	2 I48745	semaphorin B - mou
18	356.5	6.5	653	2 T03102	semaphorin homolog
19	322.5	5.9	676	2 T13937	hypothetical prote
20	238	4.4	1945	2 T13937	plexin A - fruit f
21	217	4.0	403	2 E42521	A39R protein - vac
22	208	3.8	441	2 S29921	hypothetical prote
23	199.5	3.7	1884	2 UC4975	plexin 2 precursor
24	182	3.3	1905	2 T51553	plexin - African c
25	163.5	3.0	3968	2 A44265	trichorax homolog
26	158	2.9	2051	2 T13164	plexin B - fruit f
27	155.5	2.8	1894	2 UC4980	plexin 1 precursor
28	155	2.8	1375	1 UC5148	hepatocyte growth
29	155	2.8	2352	2 T30201	Notch homolog prot

30	153	2.8	625	2 S48941	regulatory protein
31	147.5	2.7	3507	2 T34513	hypothetical prote
32	147	2.7	1425	2 T30811	hepatocyte growth
33	146.5	2.7	295	2 UC1775	salivary protein - v
34	140	2.6	1390	1 TVH0ME	hepatocyte growth
35	139.5	2.6	3869	2 A48205	All-1 protein +GTE
36	138.5	2.5	1065	2 S19482	hypothetical prote
37	138	2.5	728	2 S48569	hypothetical zinc
38	137	2.5	867	2 T41308	hypothetical poly
39	137	2.5	2492	1 C44213	nonstructural poly
40	136.5	2.5	1347	2 T02214	ubiquitous RPR mot
41	136	2.5	1310	2 T40135	oxysterol-binding
42	136	2.5	2531	2 S18188	notch protein homo
43	133.5	2.4	1829	2 T14280	RMI protein - mus
44	133.5	2.4	2187	2 T30826	nascent polypeptid
45	133.5	2.4	2531	2 A46019	notch-1 protein -

ALIGNMENTS

RESULT 1

JH0798
fasciclin IV precursor - American bird grasshopper
C:Species: Schistocerca americana (American bird grasshopper)
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: JH0798
R:Kolodkin, A.; Kachtes, D.J.; O'Connor, T.P.; Patel, N.H.; Admon, A.; Bentley, D.; Good
Neuron 9, 831-845, 1992
A>Title: Fasciclin IV: Sequence, expression, and function during growth cone guidance in
A:Reference number: JH0798; MUID:93040225; PMID:1418998
A:Accession: JH0798
A:Molecule type: mRNA
A:Residues: 1-730 <KOL>
A:Cross-references: GB:L00709; NID:G160844; PID:G160845
A:Experimental source: embryo
C:Comment: This protein plays a role in growth cone guidance in the developing central n
C:Keywords: glycoprotein; transmembrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-730/Product: Fasciclin IV #status predicted <SIG>
F:23-627/Domain: extracellular #status predicted <EXT>
F:628-652/Domain: transmembrane #status predicted <TM>
F:653-730/Domain: intracellular #status predicted <INT>
F:44,71,163,267,360,539/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match	17.8%	Score 971.5	DB 2	Length 730
Best Local Similarity	34.4%	Pred. No. 5.4e-57		
Matches	244	Conservative	124	Mismatches 224; Indels 117; Gaps 29
QY	8	LYFTLHFAGAGFPEDSEPISTHGNVTKQYVFVGHKPRGNTQR-----HRDIDMI 61		
DB	11	LIMVALH--AAAVNDVSP-----KMYVQF-----GEERRVQRFIGNESHKDFKL 54		
QY	62	MINGTGLIYARDHIYTVDID--TSHTTEIYCKSLTKWKSROADVDTCMKKHDECGN 119		
DB	55	EKDHNLSLVGANNIYVNIISRLDTEFTQ-----RIEHHSSGAHRELCTLKESSEDCON 109		
QY	120	FKVLKNDKDALFVCGTNAFNPSCRYKMDLTFEFGD-----EESGMARCFYDAKANAV 174		
DB	110	YRVLAKTDDDDVLLICGTAAYKPLCRHYALKD---GDVVEKYEYGRGLCFDPDPHNGT 165		
QY	175	ALFADGKLYSATVTPLAIDAVIYRSLGESPTLRTVHKDSKWLKEPYVOAVDYGDIYF 234		
DB	166	AIYSEGLYSALVADPSGIDPLIYG-----PLRTSRDLKQJNAFNVMTWEYNDPIFF 220		
QY	235	PREINAVEYNTMGKVFPPVAVOYCKNDMGSGQVLEKQMTSTFLKALINCSVGDSEFFYN 294		
DB	221	FRRETAVERIYNGKALYSRVARCKDKGPPHOFGR-WTSFLKSLNCSVSDYFFYFN 279		
QY	295	IIQAVTDVIRIN-GRDV---VLATSTPNYSIPGSAVCAVMDIISVFTGRFKEQKSPD 350		
DB	280	ELQSTSDIIEGNGVGOVEKLIYGVFTTPVNSIGGSAYCAFSKMSIIESTDGPKEQETWN 339		

QY 351 STWTPVDERVPRKPRGCCAGSSSLERVATSNEFPDDTLNFIKTHPLMDAIVPSIFNRPM 410
 Db 340 SNMLAVPSLKPEPRGGQCVND-----SRLPLPVSVNFVSKSHLMDAIVAFPTREI 391
 QY 411 FLRTMRYRLTKLAVD-----TAAGPYONHTVFLSEKGIILKFLARINSGLIN----- 461
 Db 392 LIRISIQYRFTKLAIVDQVRFPGG--KADVLFITGDDGVIAL-----NSAFSDSDTV 445
 QY 462 DSLFLEMSVYVNSEKSYDGVEDKRIKMGQLDRASSSLVYAFSTCYIKYPLGRG--ERHG 519
 Db 446 DSVVIELQYLP-----PGVYKNIYVARMODGDKLIVVSDDELLAIKLRHCGSDKIT 499
 QY 520 KCKKTCIASHDPYCGWIKGGACSH--SPN---SRLTFEODIERGNTDGLGCHNSFVAL 575
 Db 500 NCRG-CVSIQDPYCAMDNVELKCTAVGSPWSAGKRRFIONISLGEHKAQGGRRPQTEIV- 557
 QY 576 NGHSSSLPSTTSDSTA-----QEGYRSRGMKDMKHLDPDSTDEPLGAVSSHNG 628
 Db 558 ---ASPVPTQPTTKSSGDFVSHIQAEFEP--IDNEIVIGVDSNVIPNTLAEINHA 610
 QY 629 DKGVIRESYLKGHDLVPLV---TL-LAI-----AVILAFVWGAVFS 666
 Db 611 GSK-----LPSSQKRLPYTAEITLIALVTSGLGALVVGFIISGLFS 652

RESULT 2

semaphorin I - fruit fly (Drosophila melanogaster) (fragment)
 C/Species: Drosophila melanogaster
 C/Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 07-May-1999
 C/Accession: B49423
 R/Kolodkin, A.L.; Matthes, D.J.; Goodman, C.S.
 Cell 75, 1389-1399, 1993
 A/Title: The Semaphorin genes encode a family of transmembrane and secreted growth cone
 A/Reference number: A49423; MUID:94094332; PMID:8269517
 A/Status: preliminary; nucleic acid sequence not shown: not compared with conceptual tra
 A/Molecule type: mRNA
 A/Residues: 1-656 <KOL>
 A/Cross-references: GB:I26082
 C/Genetics:
 A/Gene: semal
 A/Cross-references: FlyBase:FBgn0011259

Query Match 16.1%; Score 875; DB 2; Length 656;
 Best Local Similarity 35.6%; Pred. No. 1.3e-50;
 Matches 208; Conservative 95; Mismatches 193; Indels 88; Gaps 20;
 QY 114 KDCCHNFIKVLKKNDALFVCGTNAFNSCRNYKMD---TLFPGDFSGMARCPYDA 169
 Db 1 EDDQNYIRIMVWSPGRLFVCGTNSPRMCTNYIISDSNYTLA---TKGQAVCPYDP 57
 QY 170 KHANVALPADKLSATVDFLADAVIYRSLSGSPFLRTVKGDSKMLKEPYQAVDYG 229
 Db 58 RHNSTISVADNELISGIVADFGSDPIYRE-----PLQIQYSLISINAPNFSSTOG 112
 QY 230 DYIIEFFREIAYEYNTGKVVFPVAVCYKNDMGSGQVLEKQWTSFLKARLNCVPGDS 289
 Db 113 DVFVFFPELAVFELNCKAIYSRVARCYKWDKGGPHR--FNRMTSFLKSLNCSIPDY 171
 QY 290 HPEYNILQAVTVIR---INGRDVYLAFTSTPINSISGSAVCAYMDLISVTTGPK 345
 Db 172 PFYFNEIOSASNLVEGQYGSWSSKLIYGVFNTPSNISIGSAVCAFDLADIDTEGGQPK 231
 QY 346 QKSPSTWTPVDERVPRKPRGCCAGSSSLERVATSNEFPDDTLNFIKTHPLMDAIVPSI 405
 Db 232 QTGINSNMLPVNNAVPRPRGSC-----HNSRRLPPTLNFITKTHSLMDENPAF 283
 QY 406 ENRPFRLTVNRYRLTKLAVD---TAAGPYONHTVFLSEKGIILKFL--ARIGNSGFL 460
 Db 284 FSQPLVATSTIYRTQIAVDAQIKTPGG--KTYDVIVFGTDHGIKISVAESADSADK 341
 QY 461 NDSLFLBMSVYVNSEKSYDGVEDKRIKMGQLDRASSSLVYAFSTCYIKYPLG 513

Db 342 VTSVIELEIDVLTGS---EPIRNLRIYRMQYDPRKGSIDDKGLIIVTSQVAIIQLH 397
 QY 514 RC--ERHGKCKKTCIASHDPYCGWIKEGAC--SHLSPN--SRLTFEODIERGNTDGLGDC 568
 Db 398 RCHNDKITSQSF--CYALQDPYCAMDKIAGKCRSHGAPRWLEBNYQVAVATQ----- 449
 QY 569 HNSFVALNGHSSSLPSTTSDSTAQEGYRSRGMKDMKHLDPDSTDEPLGAVSSHNG 628
 Db 450 ---HAACPBGKINSKDNANGEGKGFENDM---DLIDS-----RRQ 483
 QY 629 DKGVIRESYLKGHDLVPLV---TL-LAI-----AVILAFVWGAVFS 666
 Db 484 STDQELIDIDKNEPGQTSADIINAQYTVETLVAAVLAGSIFS 527

RESULT 3

semaphorin III precursor - human
 C/Species: Homo sapiens (man)
 C/Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 24-Sep-1999
 C/Accession: D49423
 R/Kolodkin, A.L.; Matthes, D.J.; Goodman, C.S.
 Cell 75, 1389-1399, 1993
 A/Title: The Semaphorin genes encode a family of transmembrane and secreted growth cone
 A/Reference number: A49423; MUID:94094332; PMID:8269517
 A/Accession: D49423
 A/Status: preliminary; nucleic acid sequence not shown
 A/Molecule type: mRNA
 A/Residues: 1-771 <KOL>
 A/Cross-references: GB:I26081; NID:G799328; PIDN:AAA65938.1; PID:G436560
 C/Genetics:
 A/Gene: GDB:SEMA1
 A/Cross-references: GDB:283448
 C/Superfamily: semaphorin

Query Match 15.6%; Score 852.5; DB 2; Length 771;
 Best Local Similarity 32.1%; Pred. No. 5.5e-49;
 Matches 208; Conservative 106; Mismatches 250; Indels 83; Gaps 20;
 QY 44 HKPGRNTQRHRLDQIMIMINGT-----LYIARDHYTV 80
 Db 22 YQNGKNNVPRKLKSYEMLESNNVITFNGLANSSYHTLDEERSRLVYGAKDHIFSPD 81
 QY 81 IDTSHEBELYCSKLTWKSQADVDCRMKKGK--KDCCHNFIKVLKKNDALFVCGTNA 139
 Db 82 L-----VNIKDPQKIWPVSYTRDECKWAGKDIKECANFIVLYARAYNTHLYACGTGA 136
 QY 140 FNPSCR-----NYKMDLLEPFGEF--SGMARCPYDAKHANVALPADKLSATVDFLA 192
 Db 137 FHIPTCYIELGHHPEDNIFLENSHPENGSGKSPYDPKLTASLIDGELYSGTAADFMG 196
 QY 193 IDAVIYRSLSGSEPTLRTVKGDSKMLKEPYVQA-----VDYGDYIIEFFREIAYVYNT 245
 Db 197 RDPALFRTGLHHHPIRTEQDSRWLNDPKFISAHLSIESDNPEDDKYFFREYALDGEH 256
 QY 246 MKGVIFPRVAVCYKNDMGSGQVLEKQWTSFLKARLNCVPG---DSHFYFNILQAVTD 301
 Db 257 SGKATNARIQIQICKNDG--HRSILNWKTFELKARLCSVPGNGIDTHF-----DELQD 310
 QY 302 VIRINGRD---VLAFTSTPYNISIPGSAVCAVMDLISVTFGRFKEQKSPSTWTPVP 357
 Db 311 VELMNFDPKPNPVYGVFTTSSNIFKGSACVMSQSVRAVFLGAPYAHNRDGPYQWVPY- 369
 QY 358 DERVPKPRPGCCAGSSSLERVATSNEFPDDTLNFIKTHPLMDAIVPSIFNRPWLRITVR 417
 Db 370 QGRVYPRPGTGP--SKTFGPGDSFKDLRDDVITPARSHPMVYVFPNNRNPYIKIDVN 428
 QY 418 YRLTKIADVDAAGYQNHVTVFLSEKGIILKFLARIGNSGFLNDSLFLBMSVYVNSEKC 477
 Db 429 YQFQIIVADVDAEDGQYDVFICTDVGTVLKVYSIKETMYDLEVLLEMTVFR----- 484
 QY 478 SYDGVEDKRIKMGQLDRASSSLVYAFSTCYIKYPLGRCERHKGCKKTCIASHDPYCGWIK 537

Db 485 -----EFTAIISAMELSTKQOOLYIGTAGVAGQLPHLHCDCIDYGRKACACCCARPPYCAM-- 537
 QY 538 EGACGSHLSPNS-RLTPEQDIERGNTDGLDPC---HNSFVALNGHS--SSLIPSTTSD 590
 Db 538 DGSAHSRYFPYAKRTTRRQDIRNG--DPLTFCSDLHHDNH---HGHSPERIIVGVENS 592
 QY 591 STAOEGESRGMLDMKHLSDPSTPLGAVSHNHQDKKGVRES 637
 Db 593 TFECSKRSORALVYWOFORNERKERDEI-RVDHIIIRTDGILLRS 638

RESULT 4

A49423
 semaphorin I precursor - beetle (Tribolium confusum)
 C/Species: Tribolium confusum
 C/Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 07-May-1999
 C/Accession: A49423
 R/Klodkin, A.L.; Matthes, D.J.; Goodman, C.S.
 Cell 75, 1389-1399, 1993
 A/Title: The Semaphorin genes encode a family of transmembrane and secreted growth cone
 A/Reference number: A49423; MUID:94094332; PMID:8269517
 A/Accession: A49423
 A/Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
 A/Molecule type: mRNA
 A/Residues: 1-711 <KOI>
 A/Cross-references: GB:I26080

Query Match 15.5%; Score 845.5; DB 2; Length 711;
 Best Local Similarity 28.9%; Pred. No. 1.4e-48;
 Matches 230; Conservative 130; Mismatches 270; Indels 165; Gaps 30;

QY 12 LHPAGAGFPDSEPI-----SISHGNTKQYVAVGHKPRRNTQRRHLDIQIMIMN- 65
 Db 12 LIALCHAMPPSSSKLTHFKSVESKSEFT-----GNATFPDH-----FVILNQ 54
 QY 66 --GTLTYAARPHITVDI--DTSHTBEIYCSKLLTWKSRQAVDTCRMKSKHKECHNPIK 122
 Db 55 DETSILVGGRRVYNLSIFDLSEK---GGRIDPSSDAHQCLILKGTDDCCQNYIR 110
 QY 123 VLLKKNDALFVCGTNAFNPSCRNKMDLEPFGD-EFSGNARCPYDAKMANVALFADGK 181
 Db 111 ILTSEPEKLVICGTNSYKPLCRTPAFEGKYIVEKEVEGIGLCPRYNEHNSISVYNGQ 170
 QY 182 LYSATVDFLAIDAVIYRSLGESPTLRVKDSKWLKPEYVQAVDYDYIYFFREIAY 241
 Db 171 LFSATVADFSGGDPILYRE-----PQRTLESLDKQLNAPFNVSVAVDYIYFFRYRTAV 225
 QY 242 EYNTMGKVPFVAVQVCKNDMGSGQVLEKQWTSFLKARLNCVSGDSHFYFNILQAVTD 301
 Db 226 EYMGCKVITYRVAVCCKDQKGRHQSRDR-WTSFLKARLNCSTIGBEYFFDEIGSTSD 284
 QY 302 VI--RINGRD--VVLATFSTPYNSIPGSAVCAVMDIASVFTGRFKQKSPDSTWTPV 357
 Db 285 IVEGYNSSDDEKLIIGLITTPNALGSAICAYQVADILRVEGSGFKQETINSKMLPVP 344
 QY 368 DERVXKPPGCCAGSSSERATSNFPPDPLNFKTRPLMDEAVPSIFENPFWLRTVVR 417
 Db 345 QNLVEPRPGQCVRSRI-----LPDKVNFIKTSLMED--VPALFGKRVLRVSLQ 395
 QY 418 YRLTIADVTAAGPYQNH--TVVPLGSEKGIILKFLARIGNSGFLNDSLFLIEMSVNSE 475
 Db 366 YRFTLITVDPVKITNNQYLDVLYIGTDDGKYLK----- 429
 QY 476 KQSYDVEDKRLMGNQDLRASSSLVYAFSTCY-----IKVP----- 511
 Db 430 -----AVNIPKHAHALYRKYRTSVHPGAPVKQKLIAPGYKVVVVVAGDEIR 478
 QY 512 ---LGRCEHKGCKTCTIASRPYGMKEGACGSHLSPNSLTP--EQDIENG----- 561
 Db 479 IANLHNCASKTRC-KDQVELQDPHCAMDAKQNLCVSITVTSYRPLIDVNVAGDNNKWS 537
 QY 562 --TDGLGCHNSFVALNGHSSILPSTTSDTAQEGESRGMLDMKHLSDPSTDPL 619

Db 538 PQDKDKTVLKNK-----PSEVENEIR-----NSIDEKD-----DSSDPL 572
 QY 620 GAVSHNHQDKKGVRESITLK---HDQLYPVV--LLATAVILAFWGA---VFSGITVY 671
 Db 573 IKTGLDDSDCDPV--SENSIGCAVRQQLVITYAGTLHYVVVAVSIVGLFSWYISGLSVF 631
 QY 672 CV--CDHRADVAVQKRE--KELTSPRSGMS--SVTKSLGFGDTQSKPKREAILTPL 726
 Db 632 AKHSDSQPEAPFIOHNLHERLSANQGYLTPRANKAVNLVYKSSSFRPKQMLDV 691
 QY 727 MHNGKLATPGNTAKM 741
 Db 692 SKDLNIAISDGTLOKI 706

RESULT 5

A18747
 semaphorin D - mouse
 C/Species: Mus musculus (house mouse)
 C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Sep-1999
 C/Accession: I48747
 R/Puschel, A.W.; Adams, R.H.; Betz, H.
 Neuron 14, 941-948, 1995
 A/Title: Murine semaphorin D/collapsin is a member of a diverse gene family and creates c
 A/Reference number: I48744; MUID:95267431; PMID:7748561
 A/Accession: I48747
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-772 <RES>
 A/Cross-references: EMBL:X85993; NID:9854329; PID:CAA59985.1; PID:9854330
 A/Genetics:
 A/Gene: semD
 C/Superfamily: semaphorin

Query Match 15.5%; Score 842.5; DB 2; Length 772;
 Best Local Similarity 34.3%; Pred. No. 2.6e-48;
 Matches 200; Conservative 85; Mismatches 225; Indels 73; Gaps 16;

QY 47 GRNTQCHRLDIQIMIMNGT-----LYIAARDHITVDIDT 83
 Db 25 GKNNVPRLKLSYKEMLESNNVITNGLANSSYHTFLIDBERSKLYGAKKHITSFNL-- 82
 QY 84 SHTBEIYCSKLLTWKSRQAVDTCRMKQH-KQCHNFIKYLKKNDDALFVCGTNAENP 142
 Db 83 ---VNIDPQKIWPVSVYTRRDECKMAGKQIKSCAFIKYLEAVNQHLYVACGTGAFHP 139
 QY 143 SCR-----NYKMDLPEPGBF--SGNARCPYDAKMANVALFADGKLSATVDFLAIDA 195
 Db 140 ICTYIEVGHHPEDNIFKLDOSHFPENGRGKSPYDKLTLASLLIDGELYSGTAADFMRGR 199
 QY 196 VIYSLGESPTLRVKDSKWLKPEYVQV-----VDYGDYIYFFREIAYEVNMTGK 248
 Db 200 AIFPTLGDHPIRTEQDHSKRLNDPRFISHLIPESNDPEDDKYFFPRENAIGSEHSGK 259
 QY 249 VVFPFVAVQVCKNDMGSGQVLEKQWTSFLKARLNCVSG-----DSHFYFNILQAVTDVIR 304
 Db 260 ATHARIQICKNDRG--HRSLVNKMWTFELKARLICVGPGRGIDTHF-----DELQDVF 313
 QY 305 INGRD---VVLATFSTPYNSIPGSAVCAVMDIASVFTGRFKQKSPDSTWTPVPER 360
 Db 314 MNSXDPKPIPVYGVFTTSSNIFKGSVAVCMSSMDVRVRFQVYAHROGPNQWVPY--QGR 372
 QY 361 VPKRPGCCAGSSSLERATSNFPPDPLNFKTRPLMDEAVPSIFENPFWLRTVVRYL 420
 Db 373 VVYRPPGTCR-SKTFPGFDSTKDLPDVYITGRSHPAKYNVVFPINNRPIMIKTDVNVQF 431
 QY 421 TKLIADVTAAGPYQNHTVFLGSEKGIILKFLARIGNSGFLNDSLFLIEMSVNSEKSYD 480
 Db 432 TQIVDRADYADGQVWFIGTGVTVLKVVSVPKETHNDEEVLLEMTVFR----- 484
 QY 481 GVEDKRLMGQDLDAASSSLVYAFSTCYIKVPLGCEHKGCKKCTCIASRDYCGGKKEGG 540

Db 485 --EPTTISAMELSTKQOOLYIGSTAGVQLPLHRCDIYKACAECLARDPYCAM--DGS 540

QY 541 ACSHLSPLNS-RLTFEODIERGNTDGLGCHNFFVALNGHSSSL 582

Db 541 SCSTRYFPTRKTRKTRKODIRNG--DPLTHCSDLEHDHNGHPSL 581

RESULT 6

A49069

collapsin - chicken

C:Species: Gallus gallus (chicken)

C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 24-Sep-1999

C:Accession: A49069

R:Uno, Y.; Raible, D.; Raper, J.A.

Cell 75, 217-227, 1993

A:Title: Collapsin: a protein in brain that induces the collapse and paralysis of neuron

A:Reference number: A49069; MUID:94006554; PMID:8402908

A:Accession: A49069

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-772 <UO>

A:Cross-references: GB:U02528; NID:9410078; PID:AA059638.1; PID:9410079

C:Superfamily: semaphorin

Query Match 15.4%; Score 839.5; DB 2; Length 772;

Best Local Similarity 35.2%; Pred. No. 4.1e-48;

Matches 198; Conservative 87; Mismatches 220; Indels 57; Gaps 17;

QY 68 LYIARADHIYVDIDTSHTEIYCSKLTWKSROADVTCRMKGKX-KDEGNFIKVLK 126

Db 69 LVYAGADHIFLFLQ-----VNKEYOKIYVPVSHSRDECKWAGKIDILRECANFIKVLK 123

QY 127 KNDALFVCGTNAFNFSCR-----NYKMDLEPFGEF--SGMARCPYDAKHANVALFAD 179

Db 124 VYQTHYACGTAFHFMCTYIEVSHREDNIPRMEDSHENGKRGKSPYDPLTLASLVD 183

QY 180 GKLYSATVDFLAIDAIVYRSLGESPTLTWKDSKWLKEPVVQA-----VDYGYI 232

Db 184 GELYSGTADDFMGDRDAIFRTLLGHHRIRTECHDSRWLMDPRFISAHLLPESDNPEDXI 243

QY 233 YEFRELAIEVYNNMGVVPFRVAVQCKNMGSGRVLEKQWTSFLKARLNCSPVG---D 288

Db 244 YEFRENALDGEHTGAKTCHARIGQICKNDPFG--HRLVNMKWTFLKARLNCSPVGNGID 302

QY 289 SHFYFNILQAVTDVIRINGRD---VVLATFTPYNSIPGSAVCAYMDLIDIASVFTGRK 344

Db 303 THF-----DELQVFLMNSKDPKMPYVGVFTTSSNIPKSAVCAYSMTDVRAVFLGPA 357

QY 345 EQKSPDSTWTPVDERYPKRPCCAGSSSLERYATSNRPDITLFIKTHPLMDAVALS 404

Db 358 HRDGPVQWVPY-QGRVFPYRPGTCP-SKTFGGFDSTKDLDEVIITFARSHPMYVPP 415

QY 405 IENRPFELRTWVYRLLTKLAIVDAAGYQNHVAVFVGESEKILKFLARIGNSGFNDLSI 464

Db 416 INSRPIMIKTDVYQFQIVVDVADDDGOYDMFTGTDIGYLVKAVSIPKETHLEEV 475

QY 465 FLEMSVYNSKESYDVEDKRIIMGQIDRASSSLVYAFSTCIKVPGLCEBHGCKKT 524

Db 476 LLEMYVFR-----EPTVISAMKISTKQOOLYIGSATGVSQLPLHRCSDYVAKACAE 526

QY 525 CIASRDYCGWIKGACSHLSPLNS-RLTFEODIERGNTDGLGCHNFFVALNGHSSSL 583

Db 527 CCLARDPYCAM--DGS6CSRYPFPAKRTTRKODIRNG--DPLTHCSD---LQHHDN--- 575

QY 584 PSTTSDSTAQEGYESRGMLD 605

Db 576 PSQGLEKTIYGVNSSTFLE 597

C:Date: 10-Apr-1998 #sequence_revision 08-May-1998 #text_change 17-Nov-2000

C:Accession: JC5928

R:Simmons, A.D.; Puschel, A.W.; McPherson, J.D.; Overhauser, J.; Lovett, M.

Biochem. Biophys. Res. Commun. 242, 685-691, 1998

A:Title: Molecular cloning and mapping of human semaphorin F from the Cri-du-chat candid

A:Reference number: JC5928; MUID:98125554; PMID:9464278

A:Accession: JC5928

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-1074 <SIM>

A:Cross-references: GB:U52840; NID:92772583; PID:AA09473.1; PID:92772584

A:Experimental source: brain

C:Comment: This protein disrupts normal brain development and leads to some of the featu

C:Genetics:

A:Gene: semaf

C:Superfamily: human semaphorin F; thrombospondin type 1 repeat homology

F:1-20/Domain: signal sequence #status predicted <SIG>

F:50-533/Domain: semaphorin #status predicted <SBM>

F:840-886/Domain: thrombospondin type 1 repeat homology <THR3>

F:971-993/Domain: transmembrane #status predicted <TM>

Query Match 15.2%; Score 830; DB 2; Length 1074;

Best Local Similarity 36.2%; Pred. No. 3e-47;

Matches 199; Conservative 90; Mismatches 211; Indels 50; Gaps 17;

QY 68 LYIARADHIYVDIDTSHTEIYCSKLTWKSROADVTCRMKGKXKDEGNFIKVLK 127

Db 70 LVYAGADHIFLFLQ-----EDLSIQAVEMEDDETRKACTSKSGSKSECONYIVLLV 123

QY 128 NDALFVCGTNAFNFSCRNKMDLEPFGEFSGMARCPYDAKHANVALF-ADGKLYSAT 186

Db 124 GGDRLFTCGTNAFTVTCNRSLSNAELHDOISGMAKCPYSHNSHTALTATAGELAYAT 183

QY 187 VTDFLAIDAVIYRSLGESPTLTWKDSKWLKEPVVQA-----VDYGYI 246

Db 184 AMDPFRDPAIYRSIGIIPRLTAQYNSKMLNEPVSYSYDIGNTFYEFRENAVEHD-C 242

QY 247 GKVVFPVAVQVCKNMGSGRVLEKQWTSFLKARLNCSPVGSHFYFNILQAVTDVIRIN 306

Db 243 GKTVESRAARVCKNDIGG-RFLLEDTWTTFKARLNCSPVGSHFYFNILQAVTDVIRIN 300

QY 307 GRDVVLAFTSPYNSIPGSAVCAYMDLIDIASVFTGRFQKSPDSTWTPVDERYPKRP 366

Db 301 --DLIYGFITMNVNSIASAACVFNLSAIAQFSGPFXQENSRSAWLPYPP--PNPFR 355

QY 367 GCCAGSSSLERYATSNRPDITLFIKTHPLMDAVALSIFNRPWLRTWVYRLLTKLAIV 426

Db 356 QCGTVDDGLYVNLTERNL-QDAQKFLVH---EVVQPTVTFSPFEDNSRF--SHAVD 408

QY 427 TAGPYQNHVAVFVGESEKILKFLARIGNSGFNDLSI-FLEMSVYNSKESYDVEDKR 486

Db 409 VVQGRBALVHTIYLATDYGTIKKAVVPLNGT---SSSCLEIELPEPERR-----RPF 458

QY 487 IMGQIDRASSSLVYAFSTCIKVPGLCEBHGCKKTCIASRDYCGWIKGACSHLS 546

Db 459 IRSIQTHSQSVLEFVGLREHVVKIKPLKRCQFY-RFTSTICIGADPYCGMDVNMKCTSLR 517

QY 547 ENSRLT-FEODIERGNTDGLGCHNFFVALNGHSSSLST--TTSDSTAQEGYESRGGM 603

Db 518 ELSMTQWQOSISA-----CPTNLTVDGHPGVMSWTPCTHTDGSAAV-----GSC 563

QY 604 LMKHLLDSP 613

Db 564 LCKTRCDSF 573

RESULT 7

JC5928

semaphorin F precursor - human

C:Species: Homo sapiens (man)

C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 24-Sep-1999

C:Accession: G01856

R:Sekido, Y.

submitted to the EMBL Data Library, June 1995

A:Reference number: G08634

A:Accession: G01856

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-749 <RES>

A:Cross-references: EMBL:U28369, NID:G974283, PIDN:AA00138.1, PID:G974284

C:Superfamily: semaphorin

Query Match 15.2%; Score 826; DB 2; Length 749;
Best Local Similarity 34.7%; Pred. No. 3.2e-47; Indels 56; Gaps 16;
Matches 198; Conservative 88; Mismatches 228;

```

QY 66 GTLYIARDHIYTVDDITSHTEIYCSKLTWKSROADVDTCKMKKH-KDECHNFIKVL 124
DB 65 GRFLVGENHVAISLNDISKR-----AKKLAMPAPVEMRECECNWAGDICTECNNFVRL 120
QY 125 LKKNDALFVCGTNAFNSCRNYKMD-----TLPEFGDEFSGMARCPIDAHANVAL 176
DB 121 HAYNHTLLACGTGAFHPTCAFEVGHRAEPVRLRP-GRIBDGKGSPPDPRHRAASV 179
QY 177 FADKLVSATVTDPLADAVIYRSIGESPRLRTVKHDSKMKKEPFYQAV-----DYG 229
DB 180 LVGBELYSVAADLWGRDFTIFRSIGQPSLRTBPHDSRWLNDEPKYKVFWIPESNPDD 239
QY 230 DYIYFFPREIAVE-YNTWGVFPFPAVQVCKNDMGSGQVLEKQWTSFLKARLNCVSP-- 286
DB 240 DKTYFFPREIYEAAPALGRLSVRVQICRNDLGG-QRSLVNKKWTFLLKARLVCSVPGV 298
QY 287 -GDSHFYFNILQAVTDVIRINGRD-----VLATSTPPYNSIPGSANCAVMDLDAISFTG 341
DB 299 EGDTHF-----DQLQDVFLSSRDHRTPLLYAVFSTSSIFQSGAVCYVSMNDVRAFLG 353
QY 342 RPEKCKSPDSTWTPVDEVPKRPCCAGSSSLERYATSNRPDPTLNFIKTHPLMDE 401
DB 352 GRLPHKEGPTHQWVSY-QGRVYPRPGMCP-SKTFGPSTSKDPPDVQIGRHHPLMYN 409
QY 401 AVPSIFNRPMFLRTWVRYRLTKLAVDTAAGYQNHVTFLQSEKGIILKPLA---RIGN 456
DB 410 FVLMSGRLPLFLQVAGYTFQIADRVAAADGHVDFLFGTDVGTVLKVISVKGRRPN 469
QY 457 SGFLNDSLFLERMVYNSEKSYGVDEKRLMGQJLDRASSSLVAFSTCIYKPLGCE 516
DB 470 S-----EGLLLELQVFE-----DSAAITSMQISSKQQLVYASPAVAQLALHACT 516
QY 517 RHGCKKTCIARDPYCGWIKEGACSHLSNSRLTF-EODIERGNTDGL-GDCNHS 571
DB 517 ALGRACCECIARDPYCAM--DGSACTRFQYAKRRFRQDIRNDPSTLCSGDSHS 572
QY 462 DSLFLERMVYNSEKSYGVDEKRLMGQJLDRASSSLVAFSTCIYKPLGCE 521
DB 472 ECLILEELHVFPE-----DSAAVTSMQISSKQQLVYASPAVAQLALHRAAGV 522
QY 522 KKTCTIARDPYCGWIKEGACSHLSNSRLTF-EODIERGNTDGLGDCNHSFVALNGH-- 578
DB 523 CTCECLARDPYCAM--DGYACTRFQPSAKRRFRQDIRNDPSTLCSGDSHPALLEKV 580
QY 579 -----SSSLPSTTSDSTAGCYESRGGM 603
DB 581 FGVEGSSAFLECEPRSLQARVEMTFQRAAGV 610

```

RESULT 9

148744

semaphorin A - mouse

C:Species: Mus musculus (house mouse)

C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Sep-1999

C:Accession: I48744

R:Fuschel, A.W.; Adams, R.H.; Betz, H.

Neuron 14, 941-948, 1995

A:Title: Murine semaphorin D/collyrin is a member of a diverse gene family and creates

A:Reference number: I48744, MID:95267431, PMID:7748561

A:Accession: I48744

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-748 <RES>

A:Cross-references: EMBL:X85990, NID:G854323, PIDN:CAA59982.1, PID:G854324

C:Genetics:

A:Gene: sema

C:Superfamily: semaphorin

Query Match 14.7%; Score 803; DB 2; Length 748;
Best Local Similarity 35.5%; Pred. No. 1.1e-45; Indels 62; Gaps 18;
Matches 191; Conservative 80; Mismatches 205;

```

QY 66 GTLYIARDHIYTVDDITSHTEIYCSKLTWKSROADVDTCKMKKH-KDECHNFIKVL 124
DB 65 GRFLVGENHVAISLNDISKR-----AKKLAMPAPVEMRECECNWAGDICTECNNFVRL 120
QY 125 LKKNDALFVCGTNAFNSCRNYKMDL-----EPFGDEFSGMARCPYDAKANVA 175
DB 121 HAYNHTLLACRTGAFHPTCALRWATAGTHASTGPEKLED---GKQKPYDPRHPPS 177
QY 176 LFADGKYSAITVDPLADAVIYRSIGESPRLRTVKHDSKMKKEPFYQAV-----DY 228
DB 178 LVGBELYSVAADLWGRDFTIFRSIGQPSLRTBPHDSRWLNDEPKYKVFWIPESNPDD 237
QY 229 GDIYFFPREIAVE-YNTWGVFPFPAVQVCKNDMGSGQVLEKQWTSFLKARLNCVSP-- 286
DB 238 DKTYFFPREIYEAAPALGRMSVRVQICRNDLGG-QRSLVNKKWTFLLKARLVCSVPG 296
QY 287 -GDSHFYFNILQAVTDVIRINGRD-----VLATSTPPYNSIPGSANCAVMDLDAISFT 340
DB 297 VEGDTHF-----DQLQDVFLSSRDHRTPLLYAVFSTSSIFQSGAVCYVSMNDVRAFL 351
QY 341 GRPECKSPDSTWTPVDEVPKRPCCAGSSSLERYATSNRPDPTLNFIKTHPLMDE 400
DB 352 GRLPHKEGPTHQWVSY-QGRVYPRPGMCP-SKTFGPSTSKDPPDVQIGRHHPLMYN 409
QY 401 AVPSIFNRPMFLRTWVRYRLTKLAVDTAAGYQNHVTFLQSEKGIILKPLA---RIGN 456
DB 410 FVLMSGRLPLFLQVAGYTFQIADRVAAADGHVDFLFGTDVGTVLKVISVKGRRPN 469
QY 457 SGFLNDSLFLERMVYNSEKSYGVDEKRLMGQJLDRASSSLVAFSTCIYKPLGCE 516
DB 470 S-----EGLLLELQVFE-----DSAAITSMQISSKQQLVYASPAVAQLALHACT 516
QY 517 RHGCKKTCIARDPYCGWIKEGACSHLSNSRLTF-EODIERGNTDGL-GDCNHS 571
DB 517 ALGRACCECIARDPYCAM--DGSACTRFQYAKRRFRQDIRNDPSTLCSGDSHS 572

```

RESULT 10

158169

semaphorin III - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 24-Sep-1999

C:Accession: I58169

R:Messemith, E.K.; Leonardo, E.D.; Shatz, C.J.; Tessier-Lavigne, M.; Goodman, C.S.; Ko

Neuron 14, 949-959, 1995

A:Title: Semaphorin III can function as a selective chemorepellent to pattern sensory pr

A:Reference number: I58169, MID:95267432, PMID:7748562

A:Accession: I58169

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-666 <RES>

A:Cross-references: GB:I40484, NID:G703189, PIDN:AAA73934.1, PID:G703190

C:Genetics:

A:Gene: SemaIII

C:Superfamily: semaphorin

Query Match 14.6%; Score 793.5; DB 2; Length 666;
Best Local Similarity 34.1%; Pred. No. 4e-45;
Matches 188; Conservative 89; Mismatches 226; Indels 49; Gaps 15;

```

QY 116 ECHNFIKYLKKNDALFVCGTNAFNSCR-----NYKMDTEPFEGDER--SGMARCPYD 168
DB 7 ECHNFIKYLKNDALFVCGTNAFNSCR-----NYKMDTEPFEGDER--SGMARCPYD 66
QY 169 AGHANVALFADGKYSAITVDPLADAVIYRSIGESPRLRTVKHDSKMKKEPFYQAV--- 225
DB 67 PRLTASLLIDELYSGTANWGRDFAIFTLGHHHPRITQHSRWLNDRFISAHLI 126
QY 226 ----VDYGDYIYFFPREIAVEYNTWGVFPFPAVQVCKNDMGSGQVLEKQWTSFLKARL 281

```

```

Db      127 PESDNPEDDKVYFFPRENAIDGESHGKATPHRIQICKNDFFGG-HRSIVNKKTTTLKARL 185
QY      282 NSGVG-----DSHFENILQAVTDVIRINGRD-----VVLATSTSTYNSIPGSAVCAYMUL 333
Db      186 ICSVPGPNIGIDTHF-----DELQDFVLMNSKDPKPKPIYGVFTTSSNIFKGSAAVCMSMS 240
QY      334 DIASVFTGFKFKQSPDSTWTVPDERVYKPRPGCCAGSSSLERATSNPEPDITLFIK 393
Db      241 DVRRLVLGYSARHDEGNVYMWXY-QGRVYPRPGCP-SKTGSGDSTTDLDDVITFAR 298
QY      394 THPLMDEAVPSIENRPFRLTWRVRLKIAVDTAAGYQNTVAVFLSGEKIILKFLAR 453
Db      299 SHPAMVNPVFPINNRPIMIKTDVNVQFTQIVDRAVDADGQYDVMEIGTVDTVLKVVSV 358
QY      454 IONGSLNDLSLFEEMSYVNSKCSYDGVEDKRIKMQDLRASSSLVYAFSCVYKVPFG 513
Db      359 PRETHDLEEVLEENTVPR-----EPTTISAMELSTQOQYIGSTIGAVQDLPLH 409
QY      514 RGRHCKCKKCTCIASRDPYCGWIKEGACSHLSPNS-RLTFEODIERGNTDGLGCHNSF 572
Db      410 RCDIYKACACELADPYCAM--DSSCSRYPFTKARTRRQDILRNG--DLTRGSLDQ 465
QY      573 VALNGHSSSL---LPSTTSDSTAQEGYESRGKMLDWKHLDPDSTPLGAVSHNQ 628
Db      466 HHDNHHGSLERIIYGVENSSTFLECSPEKSGRALVYQFORNEDRKEEI-XMGDHLIR 524
QY      629 DKKGVIRESYLK 640
Db      525 TEGGLRLSLQK 536

```

RESULT 11

hypothetical protein Y545B.1 - *Caenorhabditis elegans*
 C/Spectes: *Caenorhabditis elegans*
 C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C/Accession: T27165
 R:Renard, N.
 submitted to the EMBL Data Library, October 1998
 A/Reference number: Z20321
 A/Accession: T27165
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-712 <NUL>
 A/Cross-references: EMBL:AL032653; PIDN:CAA21714.1; GSPDB:GN00019; CESP:Y545B.1
 A/Experimental source: clone Y545B
 C/Genetics:
 A/Gene: CESP:Y545B.1
 A/Map position: 1
 A/Introns: 36/1; 70/2; 85/3; 194/2; 221/1; 253/3; 301/1; 352/3; 522/2; 560/1; 599/3; 655

Query Match 14.5%; Score 789; DB 2; Length 712;
 Best Local Similarity 28.7%; Pred. No. 8.8e-45;
 Matches 229; Conservative 129; Mismatches 282; Indels 158; Gaps 29;

```

Db      6  LLLFTLLHFGAGPDESEPISSHGNYTKOYF-----VFVGHKGRNTTORHRLDIQ 59
QY      7  LLLLFNV-----RSSEAIT---GGVNLARKQILNSVIGIDRFGGIGTSDSESDHF 55
Db      60 MIMIMNG-TLYIARHITVVDIDTSTHEEYSGSKULTWRSROADVDTCKMKGKHDECH 118
QY      56 KLLAADSDILVGARNVYVNLSTLST-----LSVNHKIDWKPRPAHIECIMGKSGKTDQ 110
Db      119 NFIYVLKKNDDALFVCGTNAFNPSCKNYKMDLLEPG---DEFGMARCPYDAKHANV 174
QY      111 NYIKVLARKSGVSLVCGTHAFSPKCRBY--TVTEFGIRNTRQFDQGSIPYKXNSS 167
Db      175 ALFADG--KLYSATVTDFLAIDAVIYR-SLGESEPT-----LRTVKHDSKMLKEPYVOA 225
QY      168 ALYVPGINQFLVAVITVDFVGNDAIYRKTIDETSSKSAANIRFOSYDARVLNAPNFVAI 227
Db      226 VDYGDYIYFFPRETAIVE--NTMGKVVFPFRAVQYCKMDGSGQVLEKQMTSLKARLNC 283

```

```

Db      228 FAYKEHYVWERELASPAIDNNPEQIYARVARCKXNDKGA-RPANERMTSYLKARLNC 286
QY      284 SVP-GDSHFENILQAVTDVIRI-NGRDVVLATSTSTYNSIPGSAVCAYMDIASVF-T 340
Db      287 SLPGSSSPFYNELKAVSDPIDAGNNHVVYTVSTPDSVDMGAVCKFSMKKIEEPDN 346
QY      341 GREFKQSPDSTWTVPDERVYKPRPGCCAGSSSLERATSNPEPDITLFIKTHPLMDE 400
Db      347 GTFHHQNNAAQSMMAFENEVPEKPRPGSCPDST-----KLPENTVSFLHHPILHR 398
QY      401 AVPSIENRPFRLTWRVRLKIAV---DTAAGPYQNTVAVFLSGEKIILKFLARIGNS 457
Db      399 PIPSV-AAPLIVEGADRADLTQITVLPRVAVGSH-NYDILFISGDSKVLKVEVDGNA 456
QY      458 GFLND-SLFEEMSYVNSKCSYDGVEDKRIKMQDLRASSSLVYAFSCVYKVPFGCE 516
Db      457 TVIQSATVFGQGVPLVN-----LRTKESVIVLSADEIASLPEVHCA 498
QY      517 RHGCKKCTCIASRDPYCGWIKEGACSHLSPNSRLTFEODIERGNTDGLGCHNSFVVLN 576
Db      499 QOTSCK-CVQLODPHCAMDSSIARCVHGSWTGQFIQNVVFGQSE--QCPBGITV-- 552
QY      577 GHSSSLPSTTSDSTAQEGYESRGKMLDWKHLDPDSTPLGAVSHNHQDKGVIRE 636
Db      553 -----REVFDNSESQAPEAVS-----RS 571
QY      637 SYLKGHDLVPYLLIAVILAFVNGAVPSGITVYCVCDHRKRDVAAYORKKELTHS-- 694
Db      572 GYFKHSTITVVLVAVASLISLIGA-FIGIRV-----NRWATSEPRSSASTSGSDY 625
QY      695 -----RGSMSVTKLSLFGDTQSKDPEKPAIILPLMHNKGKLTAPGTAMKILKA 745
Db      626 DSFGARALTRHDSLTTATKVDHF-----VPQSKQSVATSLVMSIVA 668
QY      746 DQHLDLVTALPPESTPT 763
Db      665 THHPKMSGHSIGINTPS 686

```

RESULT 12

semaphorin II precursor - fruit fly (*Drosophila melanogaster*)
 C/Spectes: *Drosophila melanogaster*
 C/Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 07-May-1999
 C/Accession: C49423
 R:Kolodkin, A.L.; Mathes, D.J.; Goodman, C.S.
 Cell 75, 1389-1399, 1993
 A/Title: The Semaphorin genes encode a family of transmembrane and secreted growth cone
 A/Reference number: A49423; MUID:94094312; PMID:8269517
 A/Accession: C49423
 A/Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
 A/Molecule type: mRNA
 A/Residues: 1-724 <KOL>
 A/Cross-references: GB:L26083
 C/Genetics:
 A/Gene: sema II
 A/Cross-references: FlyBase:FBgn0011260
 C/Superfamily: semaphorin

Query Match 14.3%; Score 779.5; DB 2; Length 724;
 Best Local Similarity 33.5%; Pred. No. 3.9e-44;
 Matches 193; Conservative 99; Mismatches 199; Indels 85; Gaps 22;

```

QY      33 NYTKQYVVFVGHKGRNTTORHRLD-----IQMIMN---GLTYIARHITVY 79
Db      32 NFYYERPCCTGNDQGNNNYKGHADVRENCGLYYRTFHNMEDRDTLYVGMADRVFRV 91
QY      80 DITSTHEEYSGK-KLTWRSROADVDTCKMKGKH-DECHNFIYVLLKND-DALFVCG 136
Db      92 NLQNTISSN--CNRDALNLEPTDDVYSCVSKSGQIFDCKNHVRVIGSMDOQDRLVCG 149
QY      137 TNAFNPSCKNY-----KMDTLEPFQDEFSGMARCPYDAKHANVALLPADG----- 180

```

```

Db      150 TNAANP---KQVVIYANTLHLPRSEYVIGV---LGIAKCPEDPNDNTALIVEGNNGGL 204
Qy      181 -KLVSATVDPFLAIDAVIYRS-----IGSEPTLTIVKDSWMLKEPVQAVDYG 231
Db      205 PGLYSGTNAEFTKADTVIFRTDLYNTSAKRLSEYFKFKTLKXDSWMLDKPMFVGSFDIG 264
Qy      232 IYFFFRRLAVYNNMKGVYPRVAQVCKNDMGSKORVLEKQWTSFLKARLNCSPGSHF 291
Db      265 YVFFFRRLAVYNNMGAVSRIRAVCKKQVGG-KLLAHNMAVYLARLNCSTSGEFPF 323
Qy      292 YENILOAVTVIRINGRDVILATPSTPYNS-IGSAVCAVMDLDAVFTGSKFQKSPDS 351
Db      324 YFNEIQSVYQLPSKSR--PFAITFTSTNGLSGAVGSGFHNEIQAAFNGKFKQSSNS 381
Qy      352 TWTFVDERVYKPRPGCCAGSSSLERYATSNFPDDTLNFKTHPLMDEAVPSIFNRP-W 410
Db      382 AMELVMSRVEPRPGTCVNDTS-----NLPDTVLNFRSHPLMDKAVNHNHPVY 433
Qy      411 FLRTMVRRLT-KIANDTAAGPIQNHVTVFLGSEKGIILKFLARIGNSGFLNDSLLE 468
Db      434 YKRDLVFTKLVVDXIRIDIL---NQEYIVYVYGNLRIRIKYIVQYRNGESLSKLDIFE 490
Qy      469 MSVYNSEKSYDGEDKIRIMGMDLRASSSLVYAFSTCVIKVPIGRG-ERHGGCKTCTIA 527
Db      491 VA-----PNEAIQVMEISQTRKSLYIGTDHRIKQIDLAMCKRRYDNCFR-CV- 536
Qy      528 SRDPYCGWIKEGAGCSHLSPNSRLTFEODIERGNTD 563
Db      537 -RDPYCGMDKAMNCRPY-----ELDLQDVANETS 567

```

RESULT 13

```

G02173
semaphorin III family homolog - human
C/Species: Homo sapiens (man)
C/Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 24-Sep-1999
C/Accession: G02173
R.Naylor, S.
submitted to the EMBL Data Library, October 1995
A/Reference number: G09275
A/Accession: G02173
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-753 <NAY>
A/Cross-references: EMBL:U38276; NID:g1061350; PID:AA818276.1; PID:g1061351
C/Superfamily: semaphorin

```

```

Query Match      14.1%; Score 768; DB 2; Length 753;
Best local Similarity 31.6%; Pred. No. 2,4e-43;
Matches 191; Conservative 109; Mismatches 207; Indels 98; Gaps 22;

Qy      6 LLLYTLHFAGAGPEDESEPSISGHNVTKQYPVFVGHKPGGRNTTQRHLDIQMIMT----- 61
Db      6 LLMASLLTGAMPSFTQD-----HLR---ATPRVRLSFKELKATG 43
Qy      62 -----MIMGT-----LYTAARDHIYTVIDTSHTEIYCSKLTAKSQA 102
Db      44 TAAHFNFLNTDYLILLKDEHDHDMYVGSXDYVLSLDLHDINREPLI-----IHMAAPQ 99
Qy      103 DVDITGRMGKH-KDCHNFIKVLKKNDALFVCGTNAFNPSG-----RAYKMTL 152
Db      100 RLEECVLGSKDVNGSGFVRLIQPNRTHLYVCGTGAIVNPMCTVYNGRRACQD-IYTL 158
Qy      153 EPPGDFSGMACPCPDAKHANVALFADGKLVSATVDELAIDAVIYRSIGSPFLRTYVH 212
Db      159 EFERLE-SGKGKCPDPKLDLTASALINEBLVAGVVIDFNGTDAIFRLTGGQTARTQY 217
Qy      213 DSKMLKEPVQA-----VDYGVITYFFFRRLAVYNNMGKVYPRVAQVCKNDMGSSQ 266
Db      218 NSRMLNDPSEFIHAELIPDSAEKNDKLYFFFRRSAAE-APQSPAVYARIGRIQINDGGG 276
Qy      267 RYLEKQWTSFLKARLNCSPGDS--HFYENILOAV---TVIRINGRDVILATPSTPYN 320

```

```

Db      277 CLVKN -MSTFLKARLVCSVGDGEDIGETHFDELQDVVQOQDVR---NPIYAVFTSSGS 332
Qy      321 STFGSAVCAVMDLDAVFTGSKFQKSPDSTTPVPEDEVPKPRPGCCAGSSSLERYAT 380
Db      333 VFGSAVCAVSMADIMVFEFGPAHKEGPYQMPF-SGMYPYRPDTCGGTTPSMKS 391
Qy      381 SNEFPDNLNFKTHPLMDEAVPSIFNRPWFLRTMVRRLTKXA---VDTAAGPYQNHV 437
Db      392 TKYVPEVIMFMSHPLMVAVYPLQRRPLVATGAPYRLTTLIAVDQVDSADGRYE---V 448
Qy      438 VFLGSEKGIILKFLARIGNSGFLNDSLFILEMSVYNSEKSYDGEDKIRIMGQDLRASS 497
Db      449 LFIQTRGVYQYIV-LPXDDQEMBELMLEVEYFK-----DAPVKTMTISSKQ 498
Qy      498 SLVYAFSTCVIKVPLGRCEHGGCKTCTIASRDPYCGWIKEGAGCSHLSPNS-RLTFEOD 556
Db      499 QLYVASAVGTHLSLHRCQAYGAACDCLARDPYCAM--DGAQCSRYTASSKRRSRQD 556
Qy      557 IERGN 561
Db      557 VRGN 561

```

RESULT 14

```

148748
semaphorin E - mouse
C/Species: Mus musculus (house mouse)
C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Sep-1999
C/Accession: 148748
R.Puschel, A.W.; Adams, R.H.; Betz, H.
Neuron 14, 941-948, 1995
A/Title: Murine semaphorin D/collapsin is a member of a diverse gene family and creates
A/Reference number: 148744; MUID:95267431; PMID:7748561
A/Accession: 148748
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-751 <RES>
A/Cross-references: EMBL:X85994; NID:g854331; PID:CAA59986.1; PID:g854332
A/Genetics:
A/Gene: sem
C/Superfamily: semaphorin

```

```

Query Match      13.5%; Score 737; DB 2; Length 751;
Best local Similarity 31.3%; Pred. No. 2,9e-41;
Matches 187; Conservative 103; Mismatches 224; Indels 74; Gaps 18;

Qy      24 SEPSISGHNVTKQYPVFVGHKPGGRNTTQRHLDIQMIMT--MNGTLYIARDHIYTVDI 81
Db      42 SEYSLH-----QOLDYRILMDDEDODRIYVGSKDHLSINI 79
Qy      82 DTSHTBELYCSKLTWKSROADVDTCKMKGHXQ-CHNFIKVLKKNDALFYCGTNAF 140
Db      80 NNISOEPL---SVFWPSTIKVECKWAGKDPHGGCNFRAVIGTQFRTHLYVCGSAGF 135
Qy      141 NPSC-----RAYKMTLEBFGDEF-----SGMARCEYDAKHANVALFADGKLVSATVT 188
Db      136 SPVCTIYLRGRSE-----DQVFMIDSKSGSGKRGCFNPNVATVSVMINELFSGWYI 189
Qy      189 DFLAIDAVIYRSIGSEPTLRTVVKDSKMLKEPVQA-----VDYGV-YIYFFRRLAV 241
Db      190 DFMGTDAIFRSLTFKRMQLRTDQHSKMLSEPMFVDAVIDGTDPNAKVYFFPKERT 249
Qy      242 EYNNMGKVYPRVAQVCKNDMGSGQRYVEKQWTSFLKARLNCSPGDS--HFYENILOAV 299
Db      250 DNNSTKQISHMILATICEQNDTGQ-QRSLYNNKTTTLKARLVCSYVDEDEGPEHFLBEDV 308
Qy      300 TDVIRINGR-DVLATPSTPYNS-IGSAVCAVMDLDAVFTGSKFQKSPDSTTPVVD 358
Db      309 FLLETNDNRITLVVGIFTSSSVFGSAVCYHLSDQIVENGPAHKEGPHQOLISY-Q 367
Qy      359 ERVVKPRPGCCAGSSSLERYATSNFPDDTLNFKTHPLMDEAVPSIFNRPWFLRTMVR 418

```

Db 368 GRIPYRPGTCGGAAFTNNRTTKDFPDVDVTFIRNHPIMNYSISPIHRRPLIVRIGTDY 427
 QY 419 RLTKIADVAAGYQNHHTVFLGSEKGIILKFLARIGNSGFLNDSLFLEMSVYNSKES 478
 Db 428 KYTKIADRVNADGKRVHVLFLGTDRGTQVKVVLPTNSS-ASGELLLELEVFKNH--- 483
 QY 479 YDGEDKRIWGMQIDBASSLYYAFSTCVIKVPLGRCERHKKCKTKCIASRDPYCGWIK 538
 Db 484 -----VPIITMEISSKKQQLYVSNBGSQVSLHCHITGTACADCCCLARDPYCAM--D 535
 QY 539 GCACSHSP-NSRLTFEODIERNTDGLGDCNHFVNLNGHSSSLPSTTSDSTQAE 595
 Db 536 GHCSRFYPTGKRSSRDVHRGNP--LTCGRGFNLKAYRNAETIVQYGVRRNSTPLE 591

RESULT 15

S66498
 M-sema F protein precursor - mouse
 C/Species: Mus musculus (house mouse)
 C/Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #ext_change 24-Nov-1999
 C/Accession: S66498
 R/Inagaki, S.; Furuyama, T.; Iwashashi, Y.
 FEBS Lett. 370, 269-272, 1995
 A/Title: Identification of a member of mouse semaphorin family.
 A/Reference number: S66498; PMID:95385809; PMID:7656991
 A/Accession: S66498
 A/Status: Preliminary
 A/Molecule type: mRNA
 A/Residues: 1-834 <TNA>
 A/Cross-references: EMBL:S79463; NID:g1110598; FIDN:AA835184.1; PID:g1110599
 C/Superfamily: semaphorin
 F/1-21/Domain: signal sequence #status predicted <SIG>
 F/22-834/Product: M-sema F protein #status predicted <MAT>

Query March 12.7%; Score 692; DB 2; Length 834;

Best Local Similarity 26.0%; Pred. No. 3.5e-38;

Matches 232; Conservative 131; Mismatches 326; Indels 204; Gaps 34;

QY 65 NGTLIYARDDHIYVIDIDTSHTEIYCSKLTWKSROADVDTCRMKGK-HKDECHNFIKY 123
 Db 62 SGLLYVAREKLEAFSV-----EALQLQGAISWEAPAEKKEICTQKGSNQTECFNIRF 116
 QY 124 LKKNDALFYCCGNANPNPCRYKMDTLEPFGEF-SGARCPYDAKHNVALFADGKL 182
 Db 117 LQPNSSHLYVCGYAFQPCYLYNMLFTLDRAEFEDGKCPYDPKAGHTGLVDGEL 176
 QY 183 YSATVTEFLAIDAVIYSLGSEPTLRVTKHDSKWLKEPYFVQA-----VDYGDYI 232
 Db 177 YSATLNNFLGTEPYILMYGTHHSIKT-EYIAFWLNEPHFVGSAPFVESVSGFTGDDXI 235
 QY 233 YFFPREIAVEYNTMGKVPPRAVQYCKNDWGSQVLEKOWTSLFKARLNCVPGDSHFY 292
 Db 236 YFFPREIAVEYDYDCEQVAVARVCKDMGGA-RTLQKWTTFPKARLVCASAP-DWKVY 293
 QY 293 FNIQAVTDVIRINGRDVY-LATSTPYNSIPGSANVCAYDMLDIASVFTGRFKEQKSPDS 351
 Db 294 FNOIKAVHTLKGASWHNTTFEGVFQARWGMDDLSAVCEYQLEQLQVFEQPYKEYSQAQ 353
 QY 352 TWTVPEDEKVPKPPGCGAGSSSLER-YATSNPEPDDTLNFIKTHPLMDEAVPSIFNRPW 410
 Db 354 KMAKRTDB-VSPSPGSCINWHRDNGTSSLELPDNTLNFIKTHPLMEDQVKKRLGRPL 412
 QY 411 FLRTWRYRLTKIADVTAAG-PYQNHVTVPLGSEKGIILKFLARIGNSGFLNDSL----- 464
 Db 413 LKKNTNF--THVAVADRVPLDGAITYVLFIGTGMILKAV-----SLGPMIH 459
 QY 465 FLEMSVYNSKESYDGEDKRIWGMQIDBASSLYYAFSTCVIKVPLGRCERHKKCKTK 524
 Db 460 WVEELQVFDDEP-----VESLVSGSKKVLFPAGSRQVLQSLADCTKTRFC-VD 508
 QY 525 CIASRDPYCGWIKESGAC-----SHLS--ENSRLTFEODIER----- 559
 Db 509 CVIARDPYCANVNTSRCVATTSGRSGSFLQVHVAANLDTSKMCNQYGIKKVRSIPKNTIV 568

QY 560 -GNTDGLGDCNHFVNLNGH-----SSSLPSTTSDSTQAEYSGRGMLDKHLLDSDPS 615
 Db 569 VSGTDLVLPCHLSSNLAAHMTFGSQDLPAEQPGSFLYDTGIALVW----- 616
 QY 616 TDPGLAVSSH-----HDKKG--VIRESYLKG-----HDQVPTLLAIAVILAFV 660
 Db 617 ---AAQRHSGPYRCYSEEGSTRLAASVLAAVVAGSVTLBARAPLENLGLWILAVVA 672
 QY 661 MGAVFGITVYCVCDHRRKDYAVVQKKEKELTHSRGSSSVYKLSGLFGDTQSKDPKE 720
 Db 673 LGAVCL-VLLLVLLRRR-----LRELE-----KGAAS-----ER 704
 QY 721 ALLPLMENGKLTATGNTAKMLIKADQHLLDLTALPTPESTPTLQCKRPSRSGRPM--- 777
 Db 705 TLVYPL-----ELPKEDASPPF--RPGPETDEKLMDEPV 735
 QY 778 -----ERNONLINACTKQMPWMSGPVIPNDLPLRASPSHIFSVVLPITQCGYQHEYVDQ 832
 Db 736 GYYSDGSLIKIVPGHARCPGPGSPSPPGITQGPLPS--PTRLHGGGRNSNANQYR- 792
 QY 833 PKMSEVAQVALBDQAATLEKTIKEHLSSKSPNHGVNLVENLDSLPPKVPQRE 885
 Db 793 -----LQVGEDRGS-----GHPLBELADELRKRLQORQ 822

Search completed: October 23, 2003, 17:11:31
 Job time : 30 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: October 23, 2003, 17:10:42 ; Search time 84 Seconds
(without alignments)
2053.386 Million cell updates/sec

Title: US-09-856-681-2

Perfect score: 5450
Sequence: 1 MRSALLLYFTLLHFAAGCF.....PPKPSAPLSTSMKPDNACT 1030

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 629382 seqs, 167460630 residues

Total number of hits satisfying chosen parameters: 629382

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

```
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/ECT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_NEW_PUB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5422.5	99.5	1047	US-09-957-187-85	Sequence 85, App1
2	4931	90.3	939	US-09-957-187-4	Sequence 4, App1
3	4590.5	84.2	884	US-09-957-187-6	Sequence 6, App1
4	3373	61.9	630	US-09-957-187-30	Sequence 30, App1
5	3362	61.7	626	US-09-957-187-83	Sequence 83, App1
6	2013.5	36.9	888	US-10-391-413-4	Sequence 4, App1
7	2009.5	36.9	888	US-09-931-836-35	Sequence 35, App1
8	2009.5	36.9	888	US-10-035-977-35	Sequence 35, App1
9	2009.5	36.9	888	US-10-137-870-544	Sequence 544, App
10	2009.5	36.9	888	US-10-140-018-544	Sequence 544, App
11	2009.5	36.9	888	US-10-140-021-544	Sequence 544, App
12	2009.5	36.9	888	US-10-140-274-544	Sequence 544, App
13	2009.5	36.9	888	US-10-140-471-544	Sequence 544, App
14	2009.5	36.9	888	US-10-140-807-544	Sequence 544, App
15	2009.5	36.9	888	US-10-140-922-544	Sequence 544, App

16	2009.5	36.9	888	12	US-10-140-924-544	Sequence 544, App
17	2009.5	36.9	888	12	US-10-140-926-544	Sequence 544, App
18	2009.5	36.9	888	12	US-10-141-698-544	Sequence 544, App
19	2009.5	36.9	888	12	US-10-141-702-544	Sequence 544, App
20	2009.5	36.9	888	12	US-10-141-704-544	Sequence 544, App
21	2009.5	36.9	888	12	US-10-142-421-544	Sequence 544, App
22	2009.5	36.9	888	12	US-10-142-432-544	Sequence 544, App
23	2009.5	36.9	888	12	US-10-142-767-544	Sequence 544, App
24	2009.5	36.9	888	12	US-10-143-033-544	Sequence 544, App
25	2009.5	36.9	888	12	US-10-144-994-544	Sequence 544, App
26	2009.5	36.9	888	12	US-10-145-628-544	Sequence 544, App
27	2009.5	36.9	888	12	US-10-145-631-544	Sequence 544, App
28	2009.5	36.9	888	12	US-10-145-633-544	Sequence 544, App
29	2009.5	36.9	888	12	US-10-145-746-544	Sequence 544, App
30	2009.5	36.9	888	12	US-10-145-748-544	Sequence 544, App
31	2009.5	36.9	888	12	US-10-145-823-544	Sequence 544, App
32	2009.5	36.9	888	12	US-10-145-826-544	Sequence 544, App
33	2009.5	36.9	888	12	US-10-145-870-544	Sequence 544, App
34	2009.5	36.9	888	12	US-10-145-876-544	Sequence 544, App
35	2009.5	36.9	888	12	US-10-145-895-544	Sequence 544, App
36	2009.5	36.9	888	12	US-10-146-724-544	Sequence 544, App
37	2009.5	36.9	888	12	US-10-146-725-544	Sequence 544, App
38	2009.5	36.9	888	12	US-10-146-795-544	Sequence 544, App
39	2009.5	36.9	888	12	US-10-147-495-544	Sequence 544, App
40	2009.5	36.9	888	12	US-10-147-501-544	Sequence 544, App
41	2009.5	36.9	888	12	US-10-147-504-544	Sequence 544, App
42	2009.5	36.9	888	12	US-10-147-506-544	Sequence 544, App
43	2009.5	36.9	888	12	US-10-147-509-544	Sequence 544, App
44	2009.5	36.9	888	12	US-10-147-510-544	Sequence 544, App
45	2009.5	36.9	888	12	US-10-147-511-544	Sequence 544, App

ALIGNMENTS

```
RESULT 1
US-09-957-187-85
Sequence 85, Application US/09957187
Publication No. US2003054514A1
GENERAL INFORMATION:
APPLICANT: Shimkete, Richard A.
APPLICANT: Labochelle, William
TITLE OF INVENTION: NOVEL POLYPEPTIDES AND PROTEINS ENCODED THEREBY
FILE REFERENCE: 15966-540 CIP
CURRENT FILING DATE: 2000-09-19
PRIOR FILING DATE: 60/123,667
PRIOR APPLICATION NUMBER: 1999-03-09
PRIOR FILING DATE: 09/520,781
PRIOR APPLICATION NUMBER: 60/234,082
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: 60/233,798
PRIOR FILING DATE: 2000-09-19
PRIOR APPLICATION NUMBER: 60/174,485
PRIOR FILING DATE: 2000-01-04
NUMBER OF SEQ ID NOS: 85
SOFTWARE: Patent Ver. 2.1
SEQ ID NO 85
LENGTH: 1047
TYPE: PRT
ORGANISM: Homo sapiens
US-09-957-187-85
Query Match 99.5% Score 5422.5; DB 11; Length 1047;
Best Local Similarity 98.3%; Pred. No. 0;
Matches 1029; Conservative 0; Mismatches 1; Indels 17; Gaps 1;
QY 1 MRSALLLYFTLLHFAAGCFPDSBPISISHGNYTKQYVFVGHKRGRTTORHRLDIOM 60
DB 1 MRSALLLYFTLLHFAAGCFPDSBPISISHGNYTKQYVFVGHKRGRTTORHRLDIOM 60
QY 61 IMINNGTYIARADHIYTVDDTIDTSHEEYKSKLTWKSROADVDTCRMKGKHKDECHNF 120
```

Db 61 IMIMNTLYIAARDHIYTDIDTSHTEETIYCSKLTWKSROADVDTCKRMKGKDECHNF 120
 QY 121 IKVLLKKNDDALFVCGTNAFNPSCRYKMDTLEPFGEDEFSGMARCPYDAKHANVALFADG 180
 Db 121 IKVLLKKNDDALFVCGTNAFNPSCRYKMDTLEPFGEDEFSGMARCPYDAKHANVALFADG 180
 QY 181 KLYSATVTDPLAIDAVIYISLGSPTLRKYKHSKWLKEPPYQAVDYGDIYFFPREIA 240
 Db 181 KLYSATVTDPLAIDAVIYISLGSPTLRKYKHSKWLKEPPYQAVDYGDIYFFPREIA 240
 QY 241 VEYNTMGKVPFPPVAQVCKNDMGSGQVLEKQMTSFLKARLNSVGDSHFYFNILQAVT 300
 Db 241 VEYNTMGKVPFPPVAQVCKNDMGSGQVLEKQMTSFLKARLNSVGDSHFYFNILQAVT 300
 QY 301 DVIRINGRDVVLATFTSPYNSIPGSAVCAVMDLDIASVFTGRFKEQKSPDSTWTPVDER 360
 Db 301 DVIRINGRDVVLATFTSPYNSIPGSAVCAVMDLDIASVFTGRFKEQKSPDSTWTPVDER 360
 QY 361 VPKRPPGCCAGSSSLERYATSNFPDDTLNFIKTHPLMDAVSIFNRPWFILTMVRYRL 420
 Db 361 VPKRPPGCCAGSSSLERYATSNFPDDTLNFIKTHPLMDAVSIFNRPWFILTMVRYRL 420
 QY 421 TKTAVDTAAGPYQNHVTVFLGSEKGIILKFLARIGNSGFLNDSLFLBEMSVNSEKSYD 480
 Db 421 TKTAVDTAAGPYQNHVTVFLGSEKGIILKFLARIGNSGFLNDSLFLBEMSVNSEKSYD 480
 QY 481 GVEDKRMGMQDLBRASSSLVYAFSTCYIKVPLGRCERHGCKKTCIASRDPYCGMIKEGG 540
 Db 481 GVEDKRMGMQDLBRASSSLVYAFSTCYIKVPLGRCERHGCKKTCIASRDPYCGMIKEGG 540
 QY 541 ACSHLSNRSRLTFEODIERGNTDGLGDCNHSFVALN-----GHSSSL 583
 Db 541 ACSHLSNRSRLTFEODIERGNTDGLGDCNHSFVALN-----GHSSSL 583
 QY 584 PSTTSTSTQOEGESRGMDMKHLLDSPDTPDLGAVSSHNDKKGVIRESYLGHD 643
 Db 584 PSTTSTSTQOEGESRGMDMKHLLDSPDTPDLGAVSSHNDKKGVIRESYLGHD 643
 QY 601 PSTTSTSTQOEGESRGMDMKHLLDSPDTPDLGAVSSHNDKKGVIRESYLGHD 660
 Db 601 PSTTSTSTQOEGESRGMDMKHLLDSPDTPDLGAVSSHNDKKGVIRESYLGHD 660
 QY 644 QLVVTLTALAIVLAFMGAVFSGITVYCVCDHRRKDVAVVORKEKELTHSRGSMSSVT 703
 Db 644 QLVVTLTALAIVLAFMGAVFSGITVYCVCDHRRKDVAVVORKEKELTHSRGSMSSVT 703
 QY 661 QLVVTLTALAIVLAFMGAVFSGITVYCVCDHRRKDVAVVORKEKELTHSRGSMSSVT 720
 Db 661 QLVVTLTALAIVLAFMGAVFSGITVYCVCDHRRKDVAVVORKEKELTHSRGSMSSVT 720
 QY 704 KLSGLFSDTOSKDXPEAILTPLMHNGKATPGNTAKMLIKADQHNLDLRLPTPESTPT 763
 Db 704 KLSGLFSDTOSKDXPEAILTPLMHNGKATPGNTAKMLIKADQHNLDLRLPTPESTPT 763
 QY 721 KLSGLFSDTOSKDXPEAILTPLMHNGKATPGNTAKMLIKADQHNLDLRLPTPESTPT 780
 Db 721 KLSGLFSDTOSKDXPEAILTPLMHNGKATPGNTAKMLIKADQHNLDLRLPTPESTPT 780
 QY 764 LQCKRKRSRGRERERONLINACTKMPMGSPVITPDLPLRASPSHISBVVVLPTIQ 823
 Db 764 LQCKRKRSRGRERERONLINACTKMPMGSPVITPDLPLRASPSHISBVVVLPTIQ 823
 QY 781 LQCKRKRSRGRERERONLINACTKMPMGSPVITPDLPLRASPSHISBVVVLPTIQ 840
 Db 781 LQCKRKRSRGRERERONLINACTKMPMGSPVITPDLPLRASPSHISBVVVLPTIQ 840
 QY 824 GYQHEVYDQPKMSFVAQWALEDOATLEKTIKELHLSKSPHNGVNLVENLDSLPKVPQ 883
 Db 824 GYQHEVYDQPKMSFVAQWALEDOATLEKTIKELHLSKSPHNGVNLVENLDSLPKVPQ 883
 QY 841 GYQHEVYDQPKMSFVAQWALEDOATLEKTIKELHLSKSPHNGVNLVENLDSLPKVPQ 900
 Db 841 GYQHEVYDQPKMSFVAQWALEDOATLEKTIKELHLSKSPHNGVNLVENLDSLPKVPQ 900
 QY 884 REASLGPFGASLSQTKSKRLEMHSSSYVDYKRSYPTNSLTRSHQATTLKRNNTSSN 943
 Db 884 REASLGPFGASLSQTKSKRLEMHSSSYVDYKRSYPTNSLTRSHQATTLKRNNTSSN 943
 QY 901 REASLGPFGASLSQTKSKRLEMHSSSYVDYKRSYPTNSLTRSHQATTLKRNNTSSN 960
 Db 901 REASLGPFGASLSQTKSKRLEMHSSSYVDYKRSYPTNSLTRSHQATTLKRNNTSSN 960
 QY 944 SSHLSRNGSPGRGNPPAPQVDSIOVHSSOPSGQAVTVSRQPSLNNYNSLTSGLKRT 1003
 Db 944 SSHLSRNGSPGRGNPPAPQVDSIOVHSSOPSGQAVTVSRQPSLNNYNSLTSGLKRT 1003
 QY 961 SSHLSRNGSPGRGNPPAPQVDSIOVHSSOPSGQAVTVSRQPSLNNYNSLTSGLKRT 1020
 Db 961 SSHLSRNGSPGRGNPPAPQVDSIOVHSSOPSGQAVTVSRQPSLNNYNSLTSGLKRT 1020
 QY 1004 PSIKPDVPPKPSFAPLSTSMKPNDACT 1030
 Db 1004 PSIKPDVPPKPSFAPLSTSMKPNDACT 1030
 QY 1021 PSIKPDVPPKPSFAPLSTSMKPNDACT 1047
 Db 1021 PSIKPDVPPKPSFAPLSTSMKPNDACT 1047

; APPLICANT: Shimkets, Richard A.
 ; APPLICANT: laRoche, William
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND PROTEINS ENCODED THEREBY
 ; FILE REFERENCE: 15966-540 CIP
 ; CURRENT APPLICATION NUMBER: US/09/957,187
 ; PRIOR FILING DATE: 2000-09-19
 ; PRIOR APPLICATION NUMBER: 60/123,667
 ; PRIOR FILING DATE: 1999-03-09
 ; PRIOR APPLICATION NUMBER: 09/520,781
 ; PRIOR FILING DATE: 2000-03-03
 ; PRIOR APPLICATION NUMBER: 60/234,082
 ; PRIOR FILING DATE: 2000-09-20
 ; PRIOR APPLICATION NUMBER: 60/233,798
 ; PRIOR FILING DATE: 2000-09-19
 ; PRIOR APPLICATION NUMBER: 60/174,485
 ; PRIOR FILING DATE: 2000-01-04
 ; NUMBER OF SEQ ID NOS: 85
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 4
 ; LENGTH: 939
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-957-187-4
 Query Match 90.3%; Score 4921; DB 11; Length 939;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 928; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MRSEALLTYFTLHFAAGFPEDSEPISTSHGNYTQYVFPVGHKRGNTTQGRHLDIOM 60
 Db 1 MRSEALLTYFTLHFAAGFPEDSEPISTSHGNYTQYVFPVGHKRGNTTQGRHLDIOM 60
 QY 61 IMIMNTLYIAARDHIYTDIDTSHTEETIYCSKLTWKSROADVDTCKRMKGKDECHNF 120
 Db 61 IMIMNTLYIAARDHIYTDIDTSHTEETIYCSKLTWKSROADVDTCKRMKGKDECHNF 120
 QY 121 IKVLLKKNDDALFVCGTNAFNPSCRYKMDTLEPFGEDEFSGMARCPYDAKHANVALFADG 180
 Db 121 IKVLLKKNDDALFVCGTNAFNPSCRYKMDTLEPFGEDEFSGMARCPYDAKHANVALFADG 180
 QY 181 KLYSATVTDPLAIDAVIYISLGSPTLRKYKHSKWLKEPPYQAVDYGDIYFFPREIA 240
 Db 181 KLYSATVTDPLAIDAVIYISLGSPTLRKYKHSKWLKEPPYQAVDYGDIYFFPREIA 240
 QY 241 VEYNTMGKVPFPPVAQVCKNDMGSGQVLEKQMTSFLKARLNSVGDSHFYFNILQAVT 300
 Db 241 VEYNTMGKVPFPPVAQVCKNDMGSGQVLEKQMTSFLKARLNSVGDSHFYFNILQAVT 300
 QY 301 DVIRINGRDVVLATFTSPYNSIPGSAVCAVMDLDIASVFTGRFKEQKSPDSTWTPVDER 360
 Db 301 DVIRINGRDVVLATFTSPYNSIPGSAVCAVMDLDIASVFTGRFKEQKSPDSTWTPVDER 360
 QY 361 VPKRPPGCCAGSSSLERYATSNFPDDTLNFIKTHPLMDAVSIFNRPWFILTMVRYRL 420
 Db 361 VPKRPPGCCAGSSSLERYATSNFPDDTLNFIKTHPLMDAVSIFNRPWFILTMVRYRL 420
 QY 421 TKTAVDTAAGPYQNHVTVFLGSEKGIILKFLARIGNSGFLNDSLFLBEMSVNSEKSYD 480
 Db 421 TKTAVDTAAGPYQNHVTVFLGSEKGIILKFLARIGNSGFLNDSLFLBEMSVNSEKSYD 480
 QY 481 GVEDKRMGMQDLBRASSSLVYAFSTCYIKVPLGRCERHGCKKTCIASRDPYCGMIKEGG 540
 Db 481 GVEDKRMGMQDLBRASSSLVYAFSTCYIKVPLGRCERHGCKKTCIASRDPYCGMIKEGG 540
 QY 541 ACSHLSNRSRLTFEODIERGNTDGLGDCNHSFVALN-----GHSSSL 583
 Db 541 ACSHLSNRSRLTFEODIERGNTDGLGDCNHSFVALN-----GHSSSL 583
 QY 584 PSTTSTSTQOEGESRGMDMKHLLDSPDTPDLGAVSSHNDKKGVIRESYLGHD 643
 Db 584 PSTTSTSTQOEGESRGMDMKHLLDSPDTPDLGAVSSHNDKKGVIRESYLGHD 643
 QY 601 PSTTSTSTQOEGESRGMDMKHLLDSPDTPDLGAVSSHNDKKGVIRESYLGHD 660
 Db 601 PSTTSTSTQOEGESRGMDMKHLLDSPDTPDLGAVSSHNDKKGVIRESYLGHD 660
 QY 644 QLVVTLTALAIVLAFMGAVFSGITVYCVCDHRRKDVAVVORKEKELTHSRGSMSSVT 703
 Db 644 QLVVTLTALAIVLAFMGAVFSGITVYCVCDHRRKDVAVVORKEKELTHSRGSMSSVT 703
 QY 661 QLVVTLTALAIVLAFMGAVFSGITVYCVCDHRRKDVAVVORKEKELTHSRGSMSSVT 720
 Db 661 QLVVTLTALAIVLAFMGAVFSGITVYCVCDHRRKDVAVVORKEKELTHSRGSMSSVT 720
 QY 704 KLSGLFSDTOSKDXPEAILTPLMHNGKATPGNTAKMLIKADQHNLDLRLPTPESTPT 763
 Db 704 KLSGLFSDTOSKDXPEAILTPLMHNGKATPGNTAKMLIKADQHNLDLRLPTPESTPT 763
 QY 721 KLSGLFSDTOSKDXPEAILTPLMHNGKATPGNTAKMLIKADQHNLDLRLPTPESTPT 780
 Db 721 KLSGLFSDTOSKDXPEAILTPLMHNGKATPGNTAKMLIKADQHNLDLRLPTPESTPT 780
 QY 764 LQCKRKRSRGRERERONLINACTKMPMGSPVITPDLPLRASPSHISBVVVLPTIQ 823
 Db 764 LQCKRKRSRGRERERONLINACTKMPMGSPVITPDLPLRASPSHISBVVVLPTIQ 823
 QY 781 LQCKRKRSRGRERERONLINACTKMPMGSPVITPDLPLRASPSHISBVVVLPTIQ 840
 Db 781 LQCKRKRSRGRERERONLINACTKMPMGSPVITPDLPLRASPSHISBVVVLPTIQ 840
 QY 824 GYQHEVYDQPKMSFVAQWALEDOATLEKTIKELHLSKSPHNGVNLVENLDSLPKVPQ 883
 Db 824 GYQHEVYDQPKMSFVAQWALEDOATLEKTIKELHLSKSPHNGVNLVENLDSLPKVPQ 883
 QY 841 GYQHEVYDQPKMSFVAQWALEDOATLEKTIKELHLSKSPHNGVNLVENLDSLPKVPQ 900
 Db 841 GYQHEVYDQPKMSFVAQWALEDOATLEKTIKELHLSKSPHNGVNLVENLDSLPKVPQ 900
 QY 884 REASLGPFGASLSQTKSKRLEMHSSSYVDYKRSYPTNSLTRSHQATTLKRNNTSSN 943
 Db 884 REASLGPFGASLSQTKSKRLEMHSSSYVDYKRSYPTNSLTRSHQATTLKRNNTSSN 943
 QY 901 REASLGPFGASLSQTKSKRLEMHSSSYVDYKRSYPTNSLTRSHQATTLKRNNTSSN 960
 Db 901 REASLGPFGASLSQTKSKRLEMHSSSYVDYKRSYPTNSLTRSHQATTLKRNNTSSN 960
 QY 944 SSHLSRNGSPGRGNPPAPQVDSIOVHSSOPSGQAVTVSRQPSLNNYNSLTSGLKRT 1003
 Db 944 SSHLSRNGSPGRGNPPAPQVDSIOVHSSOPSGQAVTVSRQPSLNNYNSLTSGLKRT 1003
 QY 961 SSHLSRNGSPGRGNPPAPQVDSIOVHSSOPSGQAVTVSRQPSLNNYNSLTSGLKRT 1020
 Db 961 SSHLSRNGSPGRGNPPAPQVDSIOVHSSOPSGQAVTVSRQPSLNNYNSLTSGLKRT 1020
 QY 1004 PSIKPDVPPKPSFAPLSTSMKPNDACT 1030
 Db 1004 PSIKPDVPPKPSFAPLSTSMKPNDACT 1030
 QY 1021 PSIKPDVPPKPSFAPLSTSMKPNDACT 1047
 Db 1021 PSIKPDVPPKPSFAPLSTSMKPNDACT 1047

Db 661 MGAFSGITTCVCDHRKRDVAVQKREKELTHSRGMSVTKLSGLFGDTOSKDPKE 720
 Qy 721 AITPLMHNGKLATPGNTAKMLIKADQHLLDLTALPTPESTPTLQOKRPSRGSEMERN 780
 Db 721 AITPLMHNGKLATPGNTAKMLIKADQHLLDLTALPTPESTPTLQOKRPSRGSEMERN 780
 Qy 781 QNLINACTKMPMPGSPVITPDLPLRASPSHPSVVLPIQOGYQHEYVQPKSEVAQ 840
 Db 781 QNLINACTKMPMPGSPVITPDLPLRASPSHPSVVLPIQOGYQHEYVQPKSEVAQ 840
 Qy 841 MALEDOATLEKYTEKITEHLSKSPNHGVLVENLDSLPKYPORASLGPASLSQTGL 900
 Db 841 MALEDOATLEKYTEKITEHLSKSPNHGVLVENLDSLPKYPORASLGPASLSQTGL 900
 Qy 901 SKRLEMHSSSYGVDYKRSYPTNSLTRSHQAT 932
 Db 901 SKRLEMHSSSYGVDYKRSYPTNSLTRSHLT 932

RESULT 3
 US-09-957-187-6
 ; Sequence 6, Application US/09957187
 ; Publication No. US20030054514A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Shinketsu, Richard A.
 ; APPLICANT: Lachochelle, William
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND PROTEINS ENCODED THEREBY
 ; FILE REFERENCE: 15966-540 CIP
 ; CURRENT APPLICATION NUMBER: US/09/957,187
 ; CURRENT FILING DATE: 2000-09-19
 ; PRIOR APPLICATION NUMBER: 60/123,667
 ; PRIOR FILING DATE: 1999-03-09
 ; PRIOR APPLICATION NUMBER: 09/520,781
 ; PRIOR FILING DATE: 2000-03-03
 ; PRIOR APPLICATION NUMBER: 60/234,082
 ; PRIOR FILING DATE: 2000-09-20
 ; PRIOR APPLICATION NUMBER: 60/233,798
 ; PRIOR FILING DATE: 2000-09-19
 ; PRIOR APPLICATION NUMBER: 60/174,485
 ; PRIOR FILING DATE: 2000-01-04
 ; NUMBER OF SEQ ID NOS: 85
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 6
 ; LENGTH: 884
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-957-187-6

Query Match 84.2%; Score 4590.5; DB 11; Length 884;
 Best Local Similarity 93.7%; Pred. No. 0;
 Matches 873; Conservative 2; Mismatches 2; Indels 55; Gaps 1;
 Qy 1 MRSSEALLTYFTLLHFAAGFPEDESEFISISHQNTYKQYVFGHKGGRNTTORHRLDIOM 60
 Db 1 MRSSEALLTYFTLLHFAAGFPEDESEFISISHQNTYKQYVFGHKGGRNTTORHRLDIOM 60
 Qy 61 IMINGGLYIARDHIYTVIDTSHTEELYSCKKLTWKSROADVDTCRMKGKHDECFANF 120
 Db 61 IMINGGLYIARDHIYTVIDTSHTEELYSCKKLTWKSROADVDTCRMKGKHDECFANF 120
 Qy 121 IKVLLKNDALFVCGTNAFNSCENYKMDTLEPFDEDESSGARCGYDAKXANVALLFADG 180
 Db 121 IKVLLKNDALFVCGTNAFNSCENYKMDTLEPFDEDESSGARCGYDAKXANVALLFADG 180
 Qy 181 KLYSATVDFLAIDATYSLGSPFLRTVKGDSKWLKEPYVOADVGYDYIFFEEREIA 240
 Db 181 KLYSATVDFLAIDATYSLGSPFLRTVKGDSKWLKEPYVOADVGYDYIFFEEREIA 240
 Qy 241 VEYNTMGKVVFPVAVQVCKNDMGSGQVLEKQWTSFLKARLNCVSYPGDSHFYFNILQAVT 300
 Db 241 VEYNTMGKVVFPVAVQVCKNDMGSGQVLEKQWTSFLKARLNCVSYPGDSHFYFNILQAVT 300

Qy 301 DVIRINGRDVLAFTSTPYNSIPGSAVCAVMDLIDIAVFGFKEQKSPDSTWTPEVDER 360
 Db 301 DVIRINGRDVLAFTSTPYNSIPGSAVCAVMDLIDIAVFGFKEQKSPDSTWTPEVDER 360
 Qy 361 VPKRPGCCAGSSGLERYATSNEDPDTLPIKTHPLMDEAVPSIENRPFRLTMVRYRL 420
 Db 361 VPKRPGCCAGSSGLERYATSNEDPDTLPIKTHPLMDEAVPSIENRPFRLTMVRYRL 420
 Qy 421 TKIAVDTAAGPYQNHVTVFGLSEKGIILKFLARIGNSGFLNDSLFLEMSVYNSEKCYD 480
 Db 421 TKIAVDTAAGPYQNHVTVFGLSEKGIILKFLARIGNSGFLNDSLFLEMSVYNSEKCYD 480
 Qy 481 GVEDKRMGMQLDRASSSLVAFSTCVIKYPLGCEHSGKCKTICIASRDYCCMIKRG 540
 Db 481 GVEDKRMGMQLDRASSSLVAFSTCVIKYPLGCEHSGKCKTICIASRDYCCMIKRG 540
 Qy 541 ACSHLSNSRLTFEODIERGNTDGLGDCNHSFVALNGHSSSLPSTTTSSTAOGYESR 600
 Db 541 ACSHLSNSRLTFEODIERGNTDGLGDCNHSFVALN----- 576
 Qy 601 GGMLDWKHLDSPESTPLGAVSHNODKKGVRESYKGDOLVPTLLAIVTLAFV 660
 Db 601 GGMLDWKHLDSPESTPLGAVSHNODKKGVRESYKGDOLVPTLLAIVTLAFV 660
 Qy 661 MGAFSGITTCVCDHRKRDVAVQKREKELTHSRGMSVTKLSGLFGDTOSKDPKE 720
 Db 661 MGAFSGITTCVCDHRKRDVAVQKREKELTHSRGMSVTKLSGLFGDTOSKDPKE 720
 Qy 721 AITPLMHNGKLATPGNTAKMLIKADQHLLDLTALPTPESTPTLQOKRPSRGSEMERN 780
 Db 721 AITPLMHNGKLATPGNTAKMLIKADQHLLDLTALPTPESTPTLQOKRPSRGSEMERN 780
 Qy 781 QNLINACTKMPMPGSPVITPDLPLRASPSHPSVVLPIQOGYQHEYVQPKSEVAQ 840
 Db 781 QNLINACTKMPMPGSPVITPDLPLRASPSHPSVVLPIQOGYQHEYVQPKSEVAQ 840
 Qy 841 MALEDOATLEKYTEKITEHLSKSPNHGVLVENLDSLPKYPORASLGPASLSQTGL 900
 Db 841 MALEDOATLEKYTEKITEHLSKSPNHGVLVENLDSLPKYPORASLGPASLSQTGL 900
 Qy 901 SKRLEMHSSSYGVDYKRSYPTNSLTRSHQAT 932
 Db 901 SKRLEMHSSSYGVDYKRSYPTNSLTRSHLT 877

RESULT 4
 US-09-957-187-30
 ; Sequence 30, Application US/09957187
 ; Publication No. US20030054514A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Shinketsu, Richard A.
 ; APPLICANT: Lachochelle, William
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND PROTEINS ENCODED THEREBY
 ; FILE REFERENCE: 15966-540 CIP
 ; CURRENT APPLICATION NUMBER: US/09/957,187
 ; CURRENT FILING DATE: 2000-09-19
 ; PRIOR APPLICATION NUMBER: 60/123,667
 ; PRIOR FILING DATE: 1999-03-09
 ; PRIOR APPLICATION NUMBER: 09/520,781
 ; PRIOR FILING DATE: 2000-03-03
 ; PRIOR APPLICATION NUMBER: 60/234,082
 ; PRIOR FILING DATE: 2000-09-20
 ; PRIOR APPLICATION NUMBER: 60/233,798
 ; PRIOR FILING DATE: 2000-09-19
 ; PRIOR APPLICATION NUMBER: 60/174,485
 ; PRIOR FILING DATE: 2000-01-04
 ; NUMBER OF SEQ ID NOS: 85
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 30
 ; LENGTH: 630
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-957-187-30

Query Match 61.9%; Score 3373; DB 11; Length 630;
 Best Local Similarity 99.8%; Pred. No. 6.3e-275;
 Matches 628; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 17 GAGFPDSEPISSHGNYTKQYVFGHKGGRNTTORHLDIQIMIMNGTLYIARDHI 76
 1 GGFPEDESPISISHGNYTKQYVFGHKGGRNTTORHLDIQIMIMNGTLYIARDHI 60
 DB 77 YVVDITSHTEIYCSKLTWKSROADVTCRMKGKDECHNFIVLLKKNDALFVCG 136
 61 YVVDITSHTEIYCSKLTWKSROADVTCRMKGKDECHNFIVLLKKNDALFVCG 120
 QY 137 TAAFNSCNRYKMDLPEFGDESGMARCPYDAKHANVALFADGKLXATVDFLAIDAV 196
 121 TAAFNSCNRYKMDLPEFGDESGMARCPYDAKHANVALFADGKLXATVDFLAIDAV 180
 DB 197 IYRSLGESPTLRVKHDSKMLKEPFYVQAVDYGDIYFFPREIAVEYNTMGKVFPRVAQ 256
 181 IYRSLGESPTLRVKHDSKMLKEPFYVQAVDYGDIYFFPREIAVEYNTMGKVFPRVAQ 240
 QY 257 VCKNDGSGQVLEKQWTSFLKARLNCVPGDSHFYFNILQAVTDVIRINGRDVLAITS 316
 241 VCKNDGSGQVLEKQWTSFLKARLNCVPGDSHFYFNILQAVTDVIRINGRDVLAITS 300
 DB 317 TPVNSTPGSAVCAVMDLIASVFTGRFKQKSPDSTWTPVDERVPRKPPGCCAGSSLE 376
 301 TPVNSTPGSAVCAVMDLIASVFTGRFKQKSPDSTWTPVDERVPRKPPGCCAGSSLE 360
 QY 377 RYATSNFPPDITNFIKTHPLMDEAVPSIFNRPMFLIRVRYLTKIAVDIAAGPYQNH 436
 361 RYATSNFPPDITNFIKTHPLMDEAVPSIFNRPMFLIRVRYLTKIAVDIAAGPYQNH 420
 DB 437 VVFLGSEKGIILKFLARIGNSGFLNDSLFLEMSVYNSEKSYDGEDKRIKMGQIDRAS 496
 421 VVFLGSEKGIILKFLARIGNSGFLNDSLFLEMSVYNSEKSYDGEDKRIKMGQIDRAS 480
 QY 497 SSIYVAFSTCVIIVPGRCEHKGKCKTCTIASRDPYCGMIKEGACSHLSPNSRLTFEED 556
 481 SSIYVAFSTCVIIVPGRCEHKGKCKTCTIASRDPYCGMIKEGACSHLSPNSRLTFEED 540
 DB 557 IERGNTDGLGDCNHSFVALNGHSSLLPSTTSDSTAQEGYESRGMLDMKHLDSPDST 616
 541 IERGNTDGLGDCNHSFVALNGHSSLLPSTTSDSTAQEGYESRGMLDMKHLDSPDST 600
 QY 617 DPLGAVSSHNDKKGVIRESYLKGHDQ 645
 601 DPLGAVSSHNDKKGVIRESYLKGHDQ 629
 DB

RESULT 5
 US-09-957-187-83
 ; Sequence 83, Application US/09957187
 ; Publication No. US200300545141
 ; GENERAL INFORMATION:
 ; APPLICANT: Shimketa, Richard A.
 ; APPLICANT: Laroche, William
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND PROTEINS ENCODED THEREBY
 ; FILE REFERENCE: 15966-340 CIP
 ; CURRENT APPLICATION NUMBER: US/09/957,187
 ; CURRENT FILING DATE: 2000-09-19
 ; PRIOR APPLICATION NUMBER: 60/123,667
 ; PRIOR FILING DATE: 1999-03-09
 ; PRIOR APPLICATION NUMBER: 09/520,781
 ; PRIOR FILING DATE: 2000-03-03
 ; PRIOR APPLICATION NUMBER: 60/234,082
 ; PRIOR FILING DATE: 2000-09-20
 ; PRIOR APPLICATION NUMBER: 60/233,798
 ; PRIOR FILING DATE: 2000-09-19
 ; PRIOR APPLICATION NUMBER: 60/114,485
 ; PRIOR FILING DATE: 2000-01-04
 ; NUMBER OF SEQ ID NOS: 85
 ; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 83
 ; LENGTH: 626
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-957-187-83

Query Match 61.7%; Score 3162; DB 11; Length 626;
 Best Local Similarity 100.0%; Pred. No. 5.2e-274;
 Matches 626; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GFPEDESPISISHGNYTKQYVFGHKGGRNTTORHLDIQIMIMNGTLYIARDHIYT 78
 1 GFPEDESPISISHGNYTKQYVFGHKGGRNTTORHLDIQIMIMNGTLYIARDHIYT 60
 DB 79 VVDITSHTEIYCSKLTWKSROADVTCRMKGKDECHNFIVLLKKNDALFVCGTN 138
 61 VVDITSHTEIYCSKLTWKSROADVTCRMKGKDECHNFIVLLKKNDALFVCGTN 120
 QY 139 AFNPSCRNRYKMDLPEFGDESGMARCPYDAKHANVALFADGKLXATVDFLAIDAV 198
 121 AFNPSCRNRYKMDLPEFGDESGMARCPYDAKHANVALFADGKLXATVDFLAIDAV 180
 DB 199 RSLGESPTLRVKHDSKMLKEPFYVQAVDYGDIYFFPREIAVEYNTMGKVFPRVAQVC 258
 181 RSLGESPTLRVKHDSKMLKEPFYVQAVDYGDIYFFPREIAVEYNTMGKVFPRVAQVC 240
 QY 259 KNDMGSGQVLEKQWTSFLKARLNCVPGDSHFYFNILQAVTDVIRINGRDVLAITS 318
 241 KNDMGSGQVLEKQWTSFLKARLNCVPGDSHFYFNILQAVTDVIRINGRDVLAITS 300
 DB 319 YNSIFGSAVCAVMDLIASVFTGRFKQKSPDSTWTPVDERVPRKPPGCCAGSSLE 378
 301 YNSIFGSAVCAVMDLIASVFTGRFKQKSPDSTWTPVDERVPRKPPGCCAGSSLE 360
 QY 379 AFSNFPDITNFIKTHPLMDEAVPSIFNRPMFLIRVRYLTKIAVDIAAGPYQNH 438
 361 AFSNFPDITNFIKTHPLMDEAVPSIFNRPMFLIRVRYLTKIAVDIAAGPYQNH 420
 DB 439 FLGSEKGIILKFLARIGNSGFLNDSLFLEMSVYNSEKSYDGEDKRIKMGQIDRAS 498
 421 FLGSEKGIILKFLARIGNSGFLNDSLFLEMSVYNSEKSYDGEDKRIKMGQIDRAS 480
 QY 499 IYVAFSTCVIIVPGRCEHKGKCKTCTIASRDPYCGMIKEGACSHLSPNSRLTFEED 558
 481 IYVAFSTCVIIVPGRCEHKGKCKTCTIASRDPYCGMIKEGACSHLSPNSRLTFEED 540
 DB 559 RGTDTGLGDCNHSFVALNGHSSLLPSTTSDSTAQEGYESRGMLDMKHLDSPDST 618
 541 RGTDTGLGDCNHSFVALNGHSSLLPSTTSDSTAQEGYESRGMLDMKHLDSPDST 600
 QY 619 LGAVSSHNDKKGVIRESYLKGHDQ 644
 601 LGAVSSHNDKKGVIRESYLKGHDQ 626
 DB

RESULT 6
 US-10-391-413-4
 ; Sequence 4, Application US/10391413
 ; Publication No. US20030167482A1
 ; GENERAL INFORMATION:
 ; APPLICANT: KIMURA, Toru et al.
 ; TITLE OF INVENTION: NOVEL SEKAPHORIN Z AND GENE ENCODING THE SAME
 ; FILE REFERENCE: 0020-5120P
 ; CURRENT APPLICATION NUMBER: US/10/391,413
 ; CURRENT FILING DATE: 2003-03-19
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 4
 ; LENGTH: 888
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-391-413-4

Query Match 36.9%; Score 2013.5; DB 12; length 888;
 Best Local Similarity 45.3%; Pred. No. 3.4e-160;
 Matches 424; Conservative 129; Mismatches 241; Indels 141; Gaps 20;

5 ALLIYFLHLHAGAGPEDESEPIISHGNTYKQYPVFVGEKKGNTTOR--HRIDIMIM 62
 12 ALLIILLILGGAGHGLFEDDEPPLISVAPRDYLNHYFVGVSGPRLTPEAGADDLNIGRVL 71
 63 IMNTQYIANDHLYTVDIDTSHREELIYCSKGLTWKSRQDVDDVDCRKKGHKQCHNFIK 122
 72 RVNRTLEFGDNDLYRVELEPPTSTELRYQKRLTRNSPNDINVCRRKSKQEGECRNFVK 131
 123 VLLKNDALFVCGTNAFNSCRNYKMDLLEPFGEDESGMARCPYDAKHAVALFADGKL 182
 132 VLLRDEBTEFLVCGSNAPNFCANYSIDTLPQVDNINSGMARCYDDKHAVALFSGML 191
 183 YSATVTEFLAIDAVIYRSLGSSPTLRIVKHDSKMLKEPYEVAQVDYGDYIYFFREIAYE 242
 192 FTATVTDPLAIDAVIYRSLGDRPTLRIVKHDSKMLKEPYEVAQVDYGDYIYFFREIAYE 251
 243 YNMGKTVFPRVAQVCKRDKGSGQVLEKQMTSFLKRLNCSVPDSEFFYNVLQAVTV 302
 252 FNYLEKVVSVARCKXNDVGSSPRVLEKQMTSFLKRLNCSVPDSEFFYNVLQAVTV 311
 303 IIRGRADVLTFTSPYNSIGSAVCAVMDIASVFTGRFEQKSPDSTWTVPYDERVP 362
 312 VSLGRFVVLAVFTSPNSIFGSAVCAFDLQVAVAFGRGRREKSPSINTVPYDEQVP 371
 363 KPRPCCAGSSSLERATSNFPPDPTNFITKHPMDFAVPSINRPFELFTMYRYRLTK 422
 372 RPRPCCAGPM--QYNASSALPDDILNFVKTHPMDFAVPSLGHAPWILRTLRHQHQLTR 429
 423 IAVDTAAPRYONHYVFLGSEKGIILKELAR--IGNSEFLNDLSFLEMSSVYNGEKGYD 480
 430 VAVDVGAPWGNQYTVFLSSAGTVLKEFLVPMNSTSGTSLGSLVEFEYRPRCRGP 489
 441 GVED--KRMGMQDLDRASSLYVAFSTCVIKVPLRCRCHKCKCTIASHRDPYCGWIK 538
 480 GGGEGQQLLSLELDAAGGLAAPRCVVAVPVARCCQYSGCMNCGISGDPYCGMAPD 549
 539 GGACSHLSNSRLTFEODIERGNTDGLDCHNSFVALNGHSSSLPSTTSDSTAQEGYE 598
 550 -GSCIFLSPGTRAAFEQDVSGASTGLGDC----- 578
 559 SRGGLMDKHLIDSDDTDLGAVSHNHQCKKGYIREYSYKQHDLVPTLLATAVLA 658
 579 -----TGLLRASLSDDRAGLVSNLVTSSVAA 606
 659 FVMGAVFSGITV-YGVCDDRDAVAVORKERE--LTSRRGSMSSVTKLSGLFQDTQSK 715
 607 FVVGAVVSGFSVGMFVGLRERELA--RRKDEALILAGAGAVLVSRL--GERAAQ 660
 716 DP-----KPEAILTPMHNGKLATPGNTAK--MLIKADQHHLDLTALPTPEST 761
 661 GPGGRRGGGGGAGVPEALLAPLMONG-----WAKATLLQGGPHDIDSLGLPTPEQT 713
 762 PTLQCKRR-----SRGSRMENONINACTKDMPMWSPVPTDLP--LBASSH 811
 714 P-LPKRFLPTPHPHALGPRAMDH-----GHPLLPASASSLLLLAPAR 757
 812 IPSVVVLTITQGYGHEYVDQPKMSEVAQMALEDOATLELYKTIENHSSKSPNGVNLV 871
 758 ABEQPPABEPTPDGLIYARGRASHGDFPLTPHASPARRVV-----SAPLGLDPA 811
 872 ENLDSLP-PKVPQREASTL-----GPPGASLSQT 898
 812 SAADGLPRWPSPPTGSLRRPLGPHAPPAATLRR 846

GENERAL INFORMATION:
 APPLICANT: Desnoyers, Inc
 APPLICANT: Eaton, Dan L.
 APPLICANT: Goddard, Audrey
 APPLICANT: Godowski, Paul J.
 APPLICANT: Gurney, Austin L.
 APPLICANT: Pan, James
 APPLICANT: Stewart, Timothy A.
 APPLICANT: Matanabe, Colin K.
 APPLICANT: Wood, William I.
 APPLICANT: Zhang, Zemin
 TITLE OR INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 FILE REFERENCE: P3030R1C1
 CURRENT APPLICATION NUMBER: US/09/931,836
 PRIOR FILING DATE: 2001-08-16
 PRIOR APPLICATION NUMBER: 60/085579
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/112514
 PRIOR FILING DATE: 1998-12-15
 PRIOR APPLICATION NUMBER: 60/113300
 PRIOR FILING DATE: 1998-12-22
 PRIOR APPLICATION NUMBER: 60/113430
 PRIOR FILING DATE: 1998-12-23
 PRIOR APPLICATION NUMBER: 60/113605
 PRIOR FILING DATE: 1998-12-23
 PRIOR APPLICATION NUMBER: 60/113621
 PRIOR FILING DATE: 1998-12-23
 PRIOR APPLICATION NUMBER: 60/114140
 PRIOR FILING DATE: 1998-12-23
 PRIOR APPLICATION NUMBER: 60/115552
 PRIOR FILING DATE: 1999-01-12
 PRIOR APPLICATION NUMBER: 60/116843
 PRIOR FILING DATE: 1999-01-22
 PRIOR APPLICATION NUMBER: 60/125774
 PRIOR FILING DATE: 1999-03-23
 PRIOR APPLICATION NUMBER: 60/125778
 PRIOR FILING DATE: 1999-03-23
 PRIOR APPLICATION NUMBER: 60/125826
 PRIOR FILING DATE: 1999-03-24
 PRIOR APPLICATION NUMBER: 60/127035
 PRIOR FILING DATE: 1999-03-31
 PRIOR APPLICATION NUMBER: 60/127706
 PRIOR FILING DATE: 1999-04-05
 PRIOR APPLICATION NUMBER: 60/129122
 PRIOR FILING DATE: 1999-04-13
 PRIOR APPLICATION NUMBER: 60/130359
 PRIOR FILING DATE: 1999-04-21
 PRIOR APPLICATION NUMBER: 60/131270
 PRIOR FILING DATE: 1999-04-27
 PRIOR APPLICATION NUMBER: 60/131272
 PRIOR FILING DATE: 1999-04-27
 PRIOR APPLICATION NUMBER: 60/131291
 PRIOR FILING DATE: 1999-04-27
 PRIOR APPLICATION NUMBER: 60/132371
 PRIOR FILING DATE: 1999-05-04
 PRIOR APPLICATION NUMBER: 60/132379
 PRIOR FILING DATE: 1999-05-04
 PRIOR APPLICATION NUMBER: 60/132383
 PRIOR FILING DATE: 1999-05-04
 PRIOR APPLICATION NUMBER: 60/135750
 PRIOR FILING DATE: 1999-05-25
 PRIOR APPLICATION NUMBER: 60/138166
 PRIOR FILING DATE: 1999-06-08
 PRIOR APPLICATION NUMBER: 60/144791
 PRIOR FILING DATE: 1999-07-20
 PRIOR APPLICATION NUMBER: 60/146370
 PRIOR FILING DATE: 1999-08-03
 PRIOR APPLICATION NUMBER: 60/162506
 PRIOR FILING DATE: 1999-10-29
 PRIOR APPLICATION NUMBER: 09/311832
 PRIOR FILING DATE: 1999-05-14
 PRIOR APPLICATION NUMBER: 09/380142

```

PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/644848
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 09/747259
PRIOR FILING DATE: 2000-12-20
PRIOR APPLICATION NUMBER: 09/816744
PRIOR FILING DATE: 2001-03-22
PRIOR APPLICATION NUMBER: 09/854208
PRIOR FILING DATE: 2001-05-10
PRIOR APPLICATION NUMBER: 09/854280
PRIOR FILING DATE: 2001-05-10
PRIOR APPLICATION NUMBER: 09/874503
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: 09/869599
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: 09/908, 827
PRIOR FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: PCT/US99/10733
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: PCT/US99/28551
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30720
PRIOR FILING DATE: 1999-12-22
PRIOR APPLICATION NUMBER: PCT/US00/05601
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: PCT/US00/05841
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: PCT/US00/14042
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: PCT/US00/15264
PRIOR FILING DATE: 2000-06-02
PRIOR APPLICATION NUMBER: PCT/US00/23522
PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: PCT/US00/23328
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: PCT/US00/32678
PRIOR FILING DATE: 2000-12-01
PRIOR APPLICATION NUMBER: PCT/US00/34956
PRIOR FILING DATE: 2000-12-20
PRIOR APPLICATION NUMBER: PCT/US01/06520
PRIOR FILING DATE: 2001-08-28
PRIOR APPLICATION NUMBER: PCT/US01/17800
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: PCT/US01/19692
PRIOR FILING DATE: 2001-06-20
PRIOR APPLICATION NUMBER: PCT/US01/21066
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: PCT/US01/21735
PRIOR FILING DATE: 2001-07-09
NUMBER OF SEQ ID NOS: 80
SEQ ID NO 35
LENGTH: 888
TYPE: PRT
ORGANISM: Homo Sapien
US-09-931-836-35

Query Match      36.9%; Score 2009.5; DB 11; Length 888;
Best Local Similarity 45.2%; Pred. No. 7.4e-160; Indels 141; Gaps 20;
Matches 423; Conservative 130; Mismatches 241;

```

```

DB 192 FTATVTDLDAIDAVIYRSLGDRPTLRVKGHSKMKPEYFVHAIVEMGSHVYFFPREIAE 251
QY 243 YNTMKAVYFPRVAQYCKNDMGSGQVLEKQMTSLFKALNCSVPDSHFEYFVILAVTV 302
DB 252 FVYLEKVVSVRAVRKNDVGSFVLEKQMTSLFKALNCSVPDSHFEYFVILAVTV 311
QY 303 IIRNGRDVLTAFSPYNSIFGSAVCAVMDLDAVFTGRFEKQSPDSTWTPVDERV 362
DB 312 VSLGRPVYLAVFSPNSIFGSAVCAVMDLDAVFTGRFEKQSPDSTWTPVDERV 371
QY 363 KRPCCGASSSLEBYATSNFPPDDTLNFKTHPLMDAVPSIENRPFELRMVRYRLTK 422
DB 372 RRPCCCAAPGW-QYNASSALPDDILNFVKTHPLMDAVPSIENRPFELRMVRYRLTK 429
QY 423 IAVDTAAGRYQNTVYFPGSEKGLIKFLAR--IGNSGNLNDFLEBMSVYNSKCSVD 480
DB 430 VAVDVAGAGWNGQTVVFGSENGVTLKFLVRENASSTGSGISVLEFEETRYPRDRCNP 489
QY 481 GVED--KIMGMQLDRASSSLYVAFSTGVIKVFLRCRCRHKCKTKCIASRDYCGWIK 538
DB 490 GGGELTGRLSLLEDAASGLLAAPRCVAVPARCQYSGCMKNCIGSDPYCGMAD 549
QY 539 GGACSHLSPNSRLTTEODIERGNTDGLDCHNSFVALNGHSSLLPSTTSSTNAEGYE 598
DB 550 -GSCIFLSPGTRAAFEODVSGASTGLGDC----- 578
QY 559 SRGMLDWKHLSDSDSTDPPLGAVSHNODKGVIRESYLKGHDQVPTLLAIAVILA 658
DB 579 -----TGLRLASLSDRGLVSVNLVTSVVA 606
QY 659 FVMGAVFSGITV-YCVCDHRRDVAVVOKEKE-LTHSRGSMSSVTKSLFGDTQSK 715
DB 607 FVMGAVFSGITV-YCVCDHRRDVAVVOKEKE-LTHSRGSMSSVTKSLFGDTQSK 715
QY 716 DP-----KPEALITPLMHNGKLAIPGTAK-MLKADQHHLDLTLPTPEST 761
DB 661 GGGGGGGGGGAGVPPENLAPLMONG-----WAKTTLQGGPHDSDGLPTPEOT 713
QY 762 PTLQCKRKP-----SRGSEMERONLINACTXDMPMGSPVPTDLP--LRASPSH 811
DB 714 P-LPQKRLPTPHPHALGPRAMDH-----GHPLLPASASSLLILAPAR 757
QY 812 IPSVVLPTTQCYGHEVDDQPMSEVMAOMLEDAATLEYTTIKHSSKPPNBNVNV 871
DB 758 AEPQPAPEPTPDGLYLAAPRASHDGFPPLTPHASPDRRRVV-----SAPTGJLDA 811
QY 872 ENUDSLP-PKVQREASL-----GPPASISQT 898
DB 812 SAADGLFRPWSPPPTGSLRRPLGPHAPPAATLRT 846

RESULT 8
US-10-035-977-35
Sequence 35, Application US/10035977
Publication No. US20030134327A1
GENERAL INFORMATION:
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Stewart, Timothy A.
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: 23030R1C10
CURRENT APPLICATION NUMBER: US/10/035,977
CURRENT FILING DATE: 2001-12-26
PRIOR APPLICATION NUMBER: 60/085579

```

PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/112514
PRIOR FILING DATE: 1998-12-15
PRIOR APPLICATION NUMBER: 60/113300
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/113430
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/113605
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/113621
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/114140
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/115552
PRIOR FILING DATE: 1998-01-12
PRIOR APPLICATION NUMBER: 60/116843
PRIOR FILING DATE: 1999-01-22
PRIOR APPLICATION NUMBER: 60/125774
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: 60/125778
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: 60/125826
PRIOR FILING DATE: 1999-03-24
PRIOR APPLICATION NUMBER: 60/127035
PRIOR FILING DATE: 1999-03-31
PRIOR APPLICATION NUMBER: 60/127706
PRIOR FILING DATE: 1999-04-05
PRIOR APPLICATION NUMBER: 60/129122
PRIOR FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: 60/130359
PRIOR FILING DATE: 1999-04-21
PRIOR APPLICATION NUMBER: 60/131270
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/131272
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/131291
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/132371
PRIOR FILING DATE: 1999-05-04
PRIOR APPLICATION NUMBER: 60/132379
PRIOR FILING DATE: 1999-05-04
PRIOR APPLICATION NUMBER: 60/133383
PRIOR FILING DATE: 1999-05-04
PRIOR APPLICATION NUMBER: 60/135750
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 60/138166
PRIOR FILING DATE: 1999-06-08
PRIOR APPLICATION NUMBER: 60/144791
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/146970
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: 60/162506
PRIOR FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: 09/311832
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 09/380142
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/644848
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 09/747259
PRIOR FILING DATE: 2000-12-20
PRIOR APPLICATION NUMBER: 09/816744
PRIOR FILING DATE: 2001-03-22
PRIOR APPLICATION NUMBER: 09/854208
PRIOR FILING DATE: 2001-05-10
PRIOR APPLICATION NUMBER: 09/854280
PRIOR FILING DATE: 2001-05-10
PRIOR APPLICATION NUMBER: 09/874503
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: 09/869599
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: 09/908,827
PRIOR FILING DATE: 2001-07-18

PRIOR APPLICATION NUMBER: PCT/US99/10733
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: PCT/US99/28551
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30720
PRIOR FILING DATE: 1999-12-22
PRIOR APPLICATION NUMBER: PCT/US00/05601
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: PCT/US00/05841
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: PCT/US00/14042
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: PCT/US00/15264
PRIOR FILING DATE: 2000-06-02
PRIOR APPLICATION NUMBER: PCT/US00/23522
PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: PCT/US00/23328
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: PCT/US00/32678
PRIOR FILING DATE: 2000-12-01
PRIOR APPLICATION NUMBER: PCT/US00/34956
PRIOR FILING DATE: 2000-12-20
PRIOR APPLICATION NUMBER: PCT/US01/06520
PRIOR FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: PCT/US01/17800
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: PCT/US01/19692
PRIOR FILING DATE: 2001-06-20
PRIOR APPLICATION NUMBER: PCT/US01/21066
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: PCT/US01/21735
PRIOR FILING DATE: 2001-07-09
NUMBER OF SEQ ID NOS: 80
SEQ ID NO 35
LENGTH: 888
TYPE: PRT
ORGANISM: Homo Sapien
US-10-035-977-35

Query Match 36.9%; Score 2009.5; DB 12; Length 886;
Best Local Similarity 45.2%; Pred. No. 7.4e-160; Indels 141; Gaps 20;

Matches 423; Conservative 130; Mismatches 241; Indels 141; Gaps 20;

QY 5 ALLVFTLLHFAAGFDESEDEISISHGNTTKQYFVVGKHKFGNTTQR--HRLDIQMIM 62
DB 12 ALLVFTLLHFAAGFDESEDEISISHGNTTKQYFVVGKHKFGNTTQR--HRLDIQMIM 62
QY 63 IMNGTLYIARADHIYTVDDITSHTEIYCSKLTWKSROADVDTCKRKKGKDECHNFIK 122
DB 72 RVNRTLEIGRDNILRYVELPPTSTELRYQKLTWRNSPNDINCRMKQGEGRNFVK 131
QY 123 VLKKNDDALFVCGTNNFNSCRYYKMDTELEPFDEESGMARCYDAKHAVALFADGL 182
DB 132 VLKKNDDALFVCGTNNFNSCRYYKMDTELEPFDEESGMARCYDAKHAVALFADGL 182
QY 132 VLKKNDDALFVCGTNNFNSCRYYKMDTELEPFDEESGMARCYDAKHAVALFADGL 182
DB 192 FTATVTDPLAIDAVIYSLGDRPFLRTVKHDSKWFKEPYFHAVERGSHVFFFEIAME 251
QY 183 YSATVTDPLAIDAVIYSLGDRPFLRTVKHDSKWFKEPYFHAVERGSHVFFFEIAME 242
DB 192 FTATVTDPLAIDAVIYSLGDRPFLRTVKHDSKWFKEPYFHAVERGSHVFFFEIAME 251
QY 243 YNTMGKVVPRVAOVCKNDKGSORVLEKQMTSTLKRNLNSVPGDSHFYNNIIQAVTDV 302
DB 252 FNYLEKTVVSRVAVCNKNDVGSFRLVEKQMTSTLKRNLNSVPGDSHFYNNIIQAVTDV 311
QY 303 IIRNGDVVLTATFSTPNNSIPGSAVCAVMDIDIAVFETGFEKQSPDSTWTVPYDERVP 362
DB 312 VSLGRPVVLAIVFTSPNSIPGSAVCAVMDIDIAVFETGFEKQSPDSTWTVPYDERVP 362
QY 363 KPRPGCCAGSSSLRBYATSNFPPDTINFITKPLMDAVALSINRPMWELTMRVRLTK 422
DB 372 RPRPGCCAGPM--QYNAASSALPDDILNFVYTHPLMEAVSLGHAAPWILRTLMRHQLTR 429
QY 423 IAVTAAQPVQNTHTVVLGSEKGIILKFLAR--IGNSGFLNDSLFLEMSYYNSKSKGYD 480

Db 430 VAVDVAGPWNQTVVFLGSEAGTVLKLVRPNASTSGTSLVLEEFETYPDRGCR 489
 QY 481 GVED--KRMGMQDLBRASSLYVAESTCYIKVPLGCEHNGKCKTKCIASRDYCGMIKE 538
 Db 490 GGGETGQRLLSTELDLAASGGLLAAPRCVVRVAVACQOQSGCMKNCISQDPYCGMAD 549
 QY 539 GGCASHLSPSRLLTFPODIERGNTDGGGCHNSFVALNGHSSSLSTTTSDTAQGYE 598
 Db 550 -GSCIFLSPGTRAFEDVSGASTSLGDC----- 578
 QY 599 SRGMLDMKHLSDPSTDPPLGAVSSHNDKKGVIRESYLKGHDQVPTLLIAIVILA 658
 Db 579 -----TGLRASLSEDRAGLVSNLLVTSVVA 606
 QY 659 FVMGAVFSGITV-YCCDHRKRDVAVVORKEKE--LTHSRGSMSSVTLSGLFGDTQSK 715
 Db 607 FVVGAVVSGFSVGMFVGLRERRELA--RRKDKKALIAHGAEGAVLSVSR-----GERRAQ 660
 QY 716 DP-----KREALITPLMNGKLTATPGNTAK-MLIKADQHLLDLTALPTPEST 761
 Db 661 GPGRGGGGGGAGVPPREALPLMONG-----WAKATLLGGPHDLSGLLPTPEQT 713
 QY 762 PTLQQRKP-----SRGSEWERNQNLINACTKMPMGSPVITDLP--LRASPSH 811
 Db 714 P-LPQKRLPTPHHPHALGPRANDH-----GHPLPASASSSLLLAIPAR 757
 QY 812 IBSVVVLPITQOQYQHRYVDQPKMSVQAOMALDQATLEYKTIKHLSSKSPNHGNNV 871
 Db 758 APQPPAPGEPPTDGLLYAARPRASHGDFPLTPHASPDRRRV-----SAFTGPLDPA 811
 QY 872 ENDSLP-PKVPQREASL-----GPPGASLSOT 898
 Db 812 SAADGLPRPWSPPPTGSLRRPLGPHADPAATLRT 846

RESULT 9

US-10-137-870-544
 ; Sequence 544, Application US/10137870
 ; Publication No. US2003013883A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: Deforge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroli, Ellen
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P33081C15
 ; CURRENT APPLICATION NUMBER: US/10/137,870
 ; PRIOR APPLICATION REMOVED - See Palm or File Wrapper
 ; NUMBER OF SEQ ID NOS: 550
 ; SEQ ID NO 544
 ; LENGTH: 888
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 ; US-10-137-870-544

Query Match 36.9%; Score 2009.5; DB 12; Length 888;
 Best Local Similarity 45.2%; Pred. No. 7,4e-160;
 Matches 423; Conservative 130; Mismatches 241; Indels 141; Gaps 20;

QY 5 ALLVFTLLHPAGAPPEDESEISISHGNYTKQYEVFGHKRGENTTOR--HRLDIQWIM 62
 Db 12 ALLLLLLLGGHGLPPEPPPLSVAPRDYLNHYEVFGSGGRLLTPAGADDLNQVLA 71
 QY 63 INNGTLYIAARDHITVVDIDTSHTEIEIYCSKLLTKRSQADVDYTCRMKGKHDECHNPIK 122
 Db 72 RYNRTLLFIQDRNLYLRYVELPEPTSTELRYQRLTKRNSPNDINVCRMKGQGECCGNFVK 131
 QY 123 VILKNDALFVCGTNAFNSCRNYKMDTLFEGFBFSMAACPYDADKANAAVALPADGL 182
 Db 132 VILLDEESTLFVCGSNAFNPVCANYSIDLQVGNISGMARCPYDPKIANVALPFDGWL 191
 QY 183 YSATYTDPLAIDAVLYRSLSGESPTLRVYHDSKWLKEPVQAVQYGDYIYFEFRIDAYE 242
 Db 192 FATATYDPLAIDAVLYRSLSGDRPTLRVYHDSKWEPEYFVAHVGSHVYFFFRIDANE 251
 QY 243 YNTMGKVPFPPVAVQYCKNDMGSGSVLEKQMSFLKARLNCVPGDSHFYFVILQAVTV 302
 Db 252 FMYLEKVVSRVARKCKNDVGSFVLEKQMSFLKARLNCVPGDSHFYFVILQAVTV 311
 QY 303 IRINGEDVLLAFSTPPYNSIPGSAYCAVMDLIDASYFTGRFQKSPDSTWTPVPERP 362
 Db 312 VSLGRFPVLAFFSTPPNSIPGSAYCAFDLTQVAAVFEGRFRPQKSPESIMTPVPERDQV 371
 QY 363 KRPDGCASSSLERATNNEFPDDTINFIKTHPLMDEAVPSIPNRPWFLRTMVRYLRK 422
 Db 372 RRPDGCAPGM-QNNASSALPDDTLNFKTHPLMDEAVPSIGHAPWILRTLMRQQLR 429
 QY 423 IAVDTAAGPYONHTVFLGSEKGIILKFLAR--IGNSGFLNLSFLEMSVYNSEKSYD 480
 Db 430 VAVDVAGPWNQTVVFLGSEAGTVLKLVRPNASTSGTSLVLEEFETYPDRGCR 489
 QY 481 GVED--KRMGMQDLBRASSLYVAESTCYIKVPLGCEHNGKCKTKCIASRDYCGMIKE 538
 Db 490 GGGETGQRLLSTELDLAASGGLLAAPRCVVRVAVACQOQSGCMKNCISQDPYCGMAD 549
 QY 539 GGCASHLSPSRLLTFPODIERGNTDGGGCHNSFVALNGHSSSLSTTTSDTAQGYE 598
 Db 550 -GSCIFLSPGTRAFEDVSGASTSLGDC----- 578
 QY 599 SRGMLDMKHLSDPSTDPPLGAVSSHNDKKGVIRESYLKGHDQVPTLLIAIVILA 658
 Db 579 -----TGLRASLSEDRAGLVSNLLVTSVVA 606
 QY 659 FVMGAVFSGITV-YCCDHRKRDVAVVORKEKE--LTHSRGSMSSVTLSGLFGDTQSK 715
 Db 607 FVVGAVVSGFSVGMFVGLRERRELA--RRKDKKALIAHGAEGAVLSVSR-----GERRAQ 660
 QY 716 DP-----KREALITPLMNGKLTATPGNTAK-MLIKADQHLLDLTALPTPEST 761
 Db 714 P-LPQKRLPTPHHPHALGPRANDH-----GHPLPASASSSLLLAIPAR 757
 QY 812 IBSVVVLPITQOQYQHRYVDQPKMSVQAOMALDQATLEYKTIKHLSSKSPNHGNNV 871
 Db 758 APQPPAPGEPPTDGLLYAARPRASHGDFPLTPHASPDRRRV-----SAFTGPLDPA 811
 QY 872 ENDSLP-PKVPQREASL-----GPPGASLSOT 898
 Db 812 SAADGLPRPWSPPPTGSLRRPLGPHADPAATLRT 846

RESULT 10

US-10-140-018-544
 ; Sequence 544, Application US/10140018
 ; Publication No. US2003013885A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen

```

APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Geriltsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Goddard, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Matanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3330R1C158
CURRENT APPLICATION NUMBER: US/10/140,018
CURRENT FILING DATE: 2002-05-06
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 544
LENGTH: 888
TYPE: PRT
ORGANISM: Homo Sapien
US-10-140-018-544

Query Match 36.9%; Score 2009.5; DB 12; Length 888;
Best Local Similarity 45.2%; Pred. No. 7.4e-160;
Matches 423; Conservative 130; Mismatches 241; Indels 141; Gaps 20;

QY 5 ALLIYFLHFAAGPEDEPISISGNATKQYPVVGVGKRPNTQR--HRLDIDMIM 62
DB 12 ALLIILLGAGLGFPEPPPLSVARPDYLNHPVVGSGPGLTPAEGADLNIQRL 71
QY 63 INNGTYIARBDHIYVDIDTSHTEIYCSKULTWKSRQADVTCRNKGKHKDCHNFIK 122
DB 72 RVRRTLFIGRDNLRYELPEPTSTELRYQKLTWNSPNDINVCNKGQEGECRFVK 131
QY 123 VLKKNDDALFVCGTNAFNSCRNYKMDLPEFGDEFSGNARCEYAKAHANVALFADGKL 182
DB 132 VLLRDBSTLFVCGSNMFPCANYSIDLQPVGDNISGMARCPYDDEKHAVALFSDGML 191
QY 183 YSAVITFLAIDATYISLSESPILRTKVDKSKLKEPPYQVADVQDIYFFPRRLAVE 242
DB 192 FTATVTDFLAIDATYISLSDGRLRTKVDKSKFKEPYFVHAWEWSHYFFPRRLAVE 251
QY 243 YNTGKVPFPPVAVQCKNDGSGORVLEKQWTSFLKRLNCSVPGDSHFYFNILQAVTDV 302
DB 252 FNYIEKVVSRVAVQCKNDGSGFRVLEKQWTSFLKRLNCSVPGDSHFYFNILQAVTV 311
QY 303 IRINGRDVVLATFSTPYNSIPGSAVCAVMDLDAVFTGRFKECKSPDSTWTPVDERVP 362
DB 312 VSLGGRPVVLAVFSTPNSIPGSAVCAFDLTQVAVAEGRFEGCKSPESIMTPVEDQVP 371
QY 363 KPRPGCCAGSSSLERYATNSPEPDDTANFKTHPLNDEAVPSINRWPLRTWRYRLTK 422
DB 372 RPRPGCCAGPM--QYNAASSALPDDILNFPVTHPLMBEAVPSLHAHWILRTLRHQLTR 429
QY 423 IAVTAAQPVONHTVVFAGSEKGIILKFLAR--IGNSGFUNDSEFLEEMSVYNEKCSYD 480
DB 430 VAVVAGAPWNGQVTVVLEGEAGTVLKLFLVAPNASTGSTSLSVLEEFYTRDRGRP 489
QY 481 GVED--KRMGMQUDRASSSLYVAFSTCVIKVPLGRGERHKGCKKTCTIASHDPYCGMIKE 538
DB 490 GGGTGTGRLLSLLEDAASGGILAFPCVAVVAVPARCQYSGCKNKGISQDPYCGWAPD 549
QY 539 GGAASHSPNRRLTFEODIERGNTDGLDCHNSVALNGHSSLLBTTTSDSTAGSGYE 598
DB 550 -GSCIFLSPGTRAFEDQVSGASTSGLDC-----578
QY 599 SRGMLDMKHLDDPSDITDPLGAVSSHNDKKGVIRESYLKGDQLQVPTLLAIVAVILA 658

```

```

DB 579 -----TGLRSLSEDRAGLVSVNLLVTSVAA 606
QY 659 FVMGAVFSGITV-YCYCDHRRKQVAVVQKKE--LTHRRGSMSSVTKLSLFGDTQSK 715
DB 607 FVVGAVVSGFVSGWPFGLRERRELA--RRKDXEAILAHGAGEVALSVSRD---GERRAQ 660
QY 716 DP-----KREALITPLMENGKLAIPGNTAK-MLIADQHHLDJLTALPPEST 761
DB 661 GEGRGGGGGGAGVPEPBLAPLMQNG-----WAKATLQGGFHDLSGLLPTEQT 713
QY 762 PTLQCKRP-----SRGSRERENQNLINACTKMDPPWGSFVITDLP---LRASPSH 811
DB 714 P-LPQKRLTPPHHPALGPAMDH-----GHPDLASASSSLLLAPAR 757
QY 812 IPSVVLPTTQCYQHEIYDQPKSEVAQMALEDQATLEYKITKHLSSKSNHGVNLY 871
DB 758 APEQPPAPGEPPTDGLYVAPRASHGDFPLTPHASPERRRVV-----SAPGTLPDPA 811
QY 872 ENIDSLP-PKVPQREASL-----GPPGASLSQT 898
DB 812 SAADGLPRPWSPPPTGSLRRLPGLPAPPAITLRT 846

RESULT 11
US-10-140-021-544
Sequence 544, Application US/10140021
Publication No. US20030138886A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Geriltsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Goddard, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Matanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3330R1C167
CURRENT APPLICATION NUMBER: US/10/140,021
CURRENT FILING DATE: 2002-05-06
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 544
LENGTH: 888
TYPE: PRT
ORGANISM: Homo Sapien
US-10-140-021-544

Query Match 36.9%; Score 2009.5; DB 12; Length 888;
Best Local Similarity 45.2%; Pred. No. 7.4e-160;
Matches 423; Conservative 130; Mismatches 241; Indels 141; Gaps 20;

QY 5 ALLIYFLHFAAGPEDEPISISGNATKQYPVVGVGKRPNTQR--HRLDIDMIM 62
DB 12 ALLIILLGAGLGFPEPPPLSVARPDYLNHPVVGSGPGLTPAEGADLNIQRL 71
QY 63 INNGTYIARBDHIYVDIDTSHTEIYCSKULTWKSRQADVTCRNKGKHKDCHNFIK 122
DB 72 RVRRTLFIGRDNLRYELPEPTSTELRYQKLTWNSPNDINVCNKGQEGECRFVK 131
QY 123 VLKKNDDALFVCGTNAFNSCRNYKMDLPEFGDEFSGNARCEYAKAHANVALFADGKL 182

```

```

Db 132 VLLRDESTLFCVGSNAEPVCANYSIDTLQFVGNISGMARCPYDPKHAVALFSDGML 191
QY 183 YSATVTDPLAIDAVIYRSLGSPRTLRTVKHDSKMLKEPYQAVDYGDIYFFFEIAYE 242
Db 192 FTAATVTDPLAIDAVIYRSLGDRPLRTVKHDSKMLKEPYQAVDYGDIYFFFEIAYE 251
QY 243 YTMKGAVFPRAVOYCKNDMGSSQVLEKQWTSFLKARLNCSPDSDSHFYENILQAVTDV 302
Db 252 FMYLEKVVSVARCKNDVGVSPVLEKQWTSFLKARLNCSPDSDSHFYENILQAVTV 311
QY 303 IIRINGADVLTATFSPYNSIPGSAYCAVMDLIDIAVFTGRFEQKSPDSTWTPVDERVP 362
Db 312 VSLGRPVVLAVFSPNSIPGSAYCAVMDLIDIAVFTGRFEQKSPDSTWTPVDERVP 371
QY 363 KRPCCGCAAGM--QYNASSALPDDILNFVKTHPLMDEAVPSIFNRPFILRTMVRJLTK 422
Db 372 RRPCCGCAAGM--QYNASSALPDDILNFVKTHPLMDEAVPSIFNRPFILRTMVRJLTK 429
QY 423 IAVDTAAGPYQNHVTVFLGSEKGIILKFLAR--IGNSGFLNDSFLEEMSVYNSKCSYD 480
Db 430 VAVDVAGAGPWGQTVFLGSEAGTVLKEFLVRNASTSGISGLVLEEFETVRPDRCGRP 489
QY 481 GVED--KRIMGQDLRASSLYVAFTCYIKVPLRCERHNGCKTKTCLASRDPYCGWIK 538
Db 490 GGGGTGQRLSLLEDAASGGLLAAPRCVYVVARCQYSGCMKNCIGSDPYCGWAPD 549
QY 539 GGACSHLSPNSRLTFEODIERGNTOGLDCHNSFVALNGHSSLLPSTTSSTAOEGYE 598
Db 550 -GSCIFLSPGTRAAFEQDVSGASTGLGDC----- 578
QY 599 SRGMLDMWHLLDSDPDSTPLGAVSSHNOCKKGVIRESYLKGHDQVPTVTLAIAVILA 658
Db 579 -----TGLRASLSERAGLVSNLLVTSVVA 606
QY 659 FVMGAVFSGITV-YCYCDHRKDVAVVQKKE--LTHSRGSMSSVTKLSGLFGDTQSK 715
Db 607 FVVGAVVSGFVGMFVGLRERELA--RRKDEKAILAHGAGVAVSVSL--GERRAQ 660
QY 716 DP-----KREALTLPLMNGKLAFTGNTAK--MLIKADQHHLDLALPTPEST 761
Db 661 GGGGAGGGGGAGVPEALPLMONG-----WAKATLLOGGPHDLSGLLPBQT 713
QY 762 PTLQQRKP-----SRGSEWERNQNLINACTKMPMGSPVPTDLP--LRASPSH 811
Db 714 P-LPQKRLPLRPHRALGPRANDH-----GHPILPASASSILLALAPAR 757
QY 812 IBSVVVLTPTQCGYQHEVYDQPMSEVAQMALEDQAATLEYKTIKHSKSPNHQVNV 871
Db 758 APQPPAPGEPPTDGLLYARPGRASHODFPILTPHASPDRRVV-----SAPTPGLDPA 811
QY 872 ENLDSLP-PKVPQREASL-----GPPGASLSQT 898
Db 812 SAADGLPRPWSPPPTGSLRRPLGPHAPPAATLART 846

```

RESULT 12

US-10-140-274-544

Sequence 544, Application US/10140274

Publication No. US20030143674A1

GENERAL INFORMATION:

```

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Denoyere, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Geritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.

```

```

APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P330R1C161
CURRENT APPLICATION NUMBER: US/10/140, 274
CURRENT FILING DATE: 2002-05-06
Prior Application removed - See file wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 544
LENGTH: 888
ORGANISM: Homo Sapien
US-10-140-274-544

```

```

Query Match 36.9%; Score 2009.5; DB 12; Length 888;
Best Local Similarity 45.2%; Pred. No. 7.4e-160; Indels 141; Gaps 20;
Matches 423; Conservative 130; Mismatches 241;

```

```

QY 5 ALLVFTLLHFAAGFPEDESEPISSHQNTYKQYVVFVGHKRGNTTOR--HRLDIQMIM 62
Db 12 ALLVFTLLHFAAGFPEDESEPISSHQNTYKQYVVFVGHKRGNTTOR--HRLDIQMIM 71
QY 63 IINGTLYAARDHYTVVDITSHTEIYCSKLLTWKSPQADVDTCKRMKKGKDECHNFIK 122
Db 72 RVNRTLFIGDRDNIYRVLELPPSTTELYQKLTWRSNPSINVCGRMKQEGECRNFK 131
QY 123 VLLKNDLALFVCGNANPFCRNKMDLTPFPGDFSGMARCPYDAKHAVALFADGKL 182
Db 132 VLLRDESTLFCVGSNAEPVCANYSIDTLQFVGNISGMARCPYDPKHAVALFSDGML 191
QY 183 YSATVTDPLAIDAVIYRSLGSPRTLRTVKHDSKMLKEPYQAVDYGDIYFFFEIAYE 242
Db 192 FTAATVTDPLAIDAVIYRSLGDRPLRTVKHDSKMLKEPYQAVDYGDIYFFFEIAYE 251
QY 243 YTMKGAVFPRAVOYCKNDMGSSQVLEKQWTSFLKARLNCSPDSDSHFYENILQAVTDV 302
Db 252 FMYLEKVVSVARCKNDVGVSPVLEKQWTSFLKARLNCSPDSDSHFYENILQAVTV 311
QY 303 IIRINGADVLTATFSPYNSIPGSAYCAVMDLIDIAVFTGRFEQKSPDSTWTPVDERVP 362
Db 312 VSLGRPVVLAVFSPNSIPGSAYCAVMDLIDIAVFTGRFEQKSPDSTWTPVDERVP 371
QY 363 KRPCCGCAAGM--QYNASSALPDDILNFVKTHPLMDEAVPSIFNRPFILRTMVRJLTK 422
Db 372 RRPCCGCAAGM--QYNASSALPDDILNFVKTHPLMDEAVPSIFNRPFILRTMVRJLTK 429
QY 423 IAVDTAAGPYQNHVTVFLGSEKGIILKFLAR--IGNSGFLNDSFLEEMSVYNSKCSYD 480
Db 430 VAVDVAGAGPWGQTVFLGSEAGTVLKEFLVRNASTSGISGLVLEEFETVRPDRCGRP 489
QY 481 GVED--KRIMGQDLRASSLYVAFTCYIKVPLRCERHNGCKTKTCLASRDPYCGWIK 538
Db 490 GGGGTGQRLSLLEDAASGGLLAAPRCVYVVARCQYSGCMKNCIGSDPYCGWAPD 549
QY 539 GGACSHLSPNSRLTFEODIERGNTOGLDCHNSFVALNGHSSLLPSTTSSTAOEGYE 598
Db 550 -GSCIFLSPGTRAAFEQDVSGASTGLGDC----- 578
QY 599 SRGMLDMWHLLDSDPDSTPLGAVSSHNOCKKGVIRESYLKGHDQVPTVTLAIAVILA 658
Db 579 -----TGLRASLSERAGLVSNLLVTSVVA 606
QY 659 FVMGAVFSGITV-YCYCDHRKDVAVVQKKE--LTHSRGSMSSVTKLSGLFGDTQSK 715
Db 607 FVVGAVVSGFVGMFVGLRERELA--RRKDEKAILAHGAGVAVSVSL--GERRAQ 660
QY 716 DP-----KREALTLPLMNGKLAFTGNTAK--MLIKADQHHLDLALPTPEST 761
Db 661 GGGGAGGGGGAGVPEALPLMONG-----WAKATLLOGGPHDLSGLLPBQT 713

```

```

QY 762 PTLQOKRRP-----SRGSEWERNONLINACTKMPGSPVITPDLR---LRASPSH 811
DB 714 P-LPQKRLPTPHHPHPLHGRAMDH-----GHPLPPLASASSLLILAPAR 757
QY 812 IPSVVLVPTIQOQYQHEHYVDQPKKSEVAQVMALEDQAATLEYKITKEHLSSKSPHGVNLV 871
DB 758 APEQPPAPGEPPTDGRLYAARPGRASHGDFPLTPHASPDRRRVY-----SAPGRLDPA 811
QY 872 ENLDSL-P-KVPGQREASL-----GPPGASLSQT 898
DB 812 SAADGLRPWSPPTGSLRRPLGPHAPPAATLRRT 846

RESULT 13
US-10-140-471-544
; Sequence 544, Application US/10140471
; Publication No. US20030138887A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P33081C163
; CURRENT APPLICATION NUMBER: US/10/140,471
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 544
; LENGTH: 888
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-471-544

Query Match 36.9%; Score 2009.5; DB 12; Length 888;
Best Local Similarity 45.2%; Pred. No. 7.4e-160; Indels 141; Gaps 20;
Matches 423; Conservative 130; Mismatches 241;

QY 5 ALLLYFTLLHFRAGFEPDESEPISSHGNYTKOYFVPGHKGRTTQR--HRLDIQIM 62
DB 12 ALLLLLLLGGAGHGFPEEPPLSVAPRDYLNHYFVFGSGPRLTPREGADLNIQVL 71
QY 63 IANGTLVIAADHLYTVUDIDTSHREITCSKLTLMKSGQAVDVCRMKSGKDKCHNRK 122
DB 72 RVNRTLFIGDRDNYRYELEPPTSTELRYQRKLTWRSNPSDINVCRMKSGKGEGRNFK 131
QY 123 VLLKKNDALFVCGTNAFNPSCRNKMDLTPFGDEFSGMARCEYDAHANVALFADGKL 182
DB 132 VLLRDESTLIVCGSNAPNVCANYSITTLQPVGNISGMARCFDPGHAVVALFSDML 191
QY 183 YSAIVTDLAIDAVIYRSLGSSPLRTVYKHSKMKIEPYVQAVDGYITFFPRELIVE 242
DB 192 FLATVTDPLAIDAVIYRSLGDRPTLRVYKHSKMKFEPYFVAHVMGSHVYFFPREIME 251
QY 243 YNTMKVVPFVAVQVCKRDKGSGQVLEKQMTSTLKAKLNSVPGDSHFYENIIQAVTDV 302
DB 252 FNYLEKVVSVARVAKNDVGGSPREVLEKQMTSTLKAKLNSVPGDSHFYENVALQAVTG 311
QY 303 IRINGRDVVLATFTSPVNSIPGSAVCAVMDLIASVFTGRFEKQKSPDSTWTVPYDERVP 362

```

```

DB 312 VSLGPRVVLAVSTSPNSIPGSAVCAFDLTQVAAVEEGFRQKSPESITWTPVEQVP 371
QY 363 KRPGCCAGSSSLERATNEPPDDLTNFKHPLMDEAVSIFNRPWFRTWRYELTK 422
DB 372 RPRPGCCAAPGM--QYNASSALPDDILNFKYHPLMDEAVPSIGHAPWILRTLMRHQLTR 429
QY 423 IAVDTAAGPYQNTHTVFLGSEKGIILKFLAR--IGNSGFLNDSLFLFEMSVNSKESYD 480
DB 430 VAVDVAGRWGNGQNTVFLGSEACTVAKFLRPAASTSGISGLVFLFEPFTYPRDRGRP 489
QY 481 GVED--KRIMGCLDRASSGLYVAFSTYKIVPLGRCERHKKCTKTIAARDPYCGMIKE 538
DB 490 GGGGTGQRLTLSTLHDAASGGLLAAPRCVYRVVPAACQYSGCKMKICISQDPCGWAPD 549
QY 539 GGCASHLSPNSRLTFPQDIERGNTDGLGCDHNSFVALNGHSSSLPPTTSDTAQGEYE 598
DB 550 -GSCIFLSPQTRAFQDVSGASTSLGDD----- 578
QY 599 SRGMLDMKHLDSPDSTDLGAVSSHNDKKGVIRESYLKGHDQVPTLLAIVILA 658
DB 579 -----TGLRASLSEDRAGIVSNLVTSSVAA 606
QY 659 FVNGAVFSGITV-YCYCDHRRKDVAVVQREKE--LTHSRGNSVYTKLSGLFGDTQSK 715
DB 607 FVVGAVVSGFSGVGMFVGLRERRELA--RRKDEALLAHGAEVLSYSRL---GERRAQ 660
QY 716 DP-----KREALTLMNGKLTATGNNAK-MLIRADQHLDTALTPTPEST 761
DB 661 GPGRGGGGGGAGVPEALALPMONG-----WKATLLQGGPHDLDSGLLPTPEQT 713
QY 762 PTLQOKRRP-----SRGSEWERNONLINACTKMPGSPVITPDLR---LRASPSH 811
DB 714 P-LPQKRLPTPHHPHPLHGRAMDH-----GHPLPPLASASSLLILAPAR 757
QY 812 IPSVVLVPTIQOQYQHEHYVDQPKKSEVAQVMALEDQAATLEYKITKEHLSSKSPHGVNLV 871
DB 758 APEQPPAPGEPPTDGRLYAARPGRASHGDFPLTPHASPDRRRVY-----SAPGRLDPA 811
QY 872 ENLDSL-P-KVPGQREASL-----GPPGASLSQT 898
DB 812 SAADGLRPWSPPTGSLRRPLGPHAPPAATLRRT 846

RESULT 14
US-10-140-807-544
; Sequence 544, Application US/10140807
; Publication No. US20030134354A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P33081C174
; CURRENT APPLICATION NUMBER: US/10/140,807
; CURRENT FILING DATE: 2002-05-07
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550

```


SEQ ID NO 544
LENGTH: 888
TYPE: PRT
ORGANISM: Homo Sapien
US-10-140-922-544

Query Match 36.9%; Score 2009.5; DB 12; Length 888;
Best Local Similarity 45.2%; Pred. No. 7.4e-160;
Matches 423; Conservative 130; Mismatches 241; Indels 141; Gaps 20;

```

5 ALLYFTLLHPAGAPPEDESEPISSHGNYTKQYVFGHKGRNTTOR--HRLDIOMIM 62
12 ALLLLLLLGGAGHGFPEEPPLSVAPRDYLNHYVFGSGGRLLTPAEGADDLNIQVLT 71
63 INNGTLYIAARDHIYTVIDTSHTEIYCSKLLTKSKROADVDTCRMKGKHECHNEFK 122
72 RYNRTFLFGDRDNLRYVELLEPPTSTELKRYKRLTWRSNPSDINVCRMKGKQEGECRNFK 131
123 VLLKNDLALFYCGTNAENPSCRNKMDLEPFGEFSGMARCPYDAKHANVALFADGKL 182
132 VLLRDESTLFCVGSNAENPVCANYSIDLQGVGNISGMARCPYDPKHANVALFSDGML 191
183 YSATYTDFLAIDAVIYRSLGSPFLRTVYKHSKWLKEPFYQAVDYGDIYFFFEIAYE 242
192 FTAIYTDFLAIDAVIYRSLGDRPTLRIVYKHSKWLKEPFYQAVDYGDIYFFFEIAYE 251
243 YNTMGKVYFPRVAQYCKNDMGSGORVLEKQWTSFLKARLNCSPGDSHFYENILQAVTV 302
252 FNYLEKVVSVARVCKNDVGSPVLEKQWTSFLKARLNCSPGDSHFYENILQAVTV 311
303 IINGRDVLAIFSTPYNSIPSSAVCAVMDLDAVFTGRFEQKSPDSTWTPVDEVRP 362
312 VSLGGRPVVLAIFSTPYNSIPSSAVCAVMDLDAVFTGRFEQKSPDSTWTPVDEVRP 371
363 KRPFGCCAGSSSLERATYNSNEPDDTLNFKTHPLMDEAVPSIFNRPMWLFRTMVRRLTK 422
372 RRPFGCCAPGM--QYNASSALPDDTLNFKTHPLMDEAVPSIFNRPMWLFRTMVRRLTK 429
423 IAVDTAGFYONHTVVFELGSEKGIILKFLAR--IGNSGFLNSLPLEMSVYNSKCSYD 480
430 VAVDVAGFWGNGTIVVFELGSEAGTVLKFVLRNASTSGTSGVLFEEBETRYRDPDRCGP 489
481 GVED--KRTMGQOLDRASSLYVASTCYIKVPLRGCEHKGCKTCLASRPYCGMIKE 538
490 GGETEQRLSLLELDAASGGLLAAPRCYVAVFVAVACQYSGCMKNCIGSDPYCGWADP 549
539 GGACSHLSPNSRLTFEODLERGTDLGDCNHSFVALNGHSSSLPSTTSDTAQEGYE 598
550 -GSCITLSEGTAAFEQDVSGASTSGLDG----- 578
599 SRGMLDMKHLDSPDSTDFLGAVSSHNDQKKGVIRESYLKGHDQLVPTLLAVALITA 658
579 -----TGLRASILEDRAVLAVNLLVTSVAA 606
659 FVWGAFVSGITV-YCYCHRRKQVAVVORKEK--LTHSRGSMSTVTLGSLFGDPTOK 715
607 FVVGAVVSGFSSVGMFWGLBERELA--RRDKKAILAHGAGEAVLVSFL---GRRKQ 660
716 DP-----KPEALITPLMENGKLTATPGNTAK--MLIKADQHHDLTALPTPEST 761
661 GPGGRGGGGGAGVPEALALPLMONG-----WAKATLLGSGHDDLSGLLPPEQT 713
762 PTLQGRKP-----SRGSRERNQNLINACTKMPNGSVITPTLP--LRASPSH 811
714 P-LPQRRLPLPHPHALGPRAMDH-----GHPLLPASASSSLLLLAIPAR 757
812 IPSVVLPTQCGYCHYVQPKMSVAVQMALEDQATLEKTIKXHLSSKSPHGVNLY 871
758 APQPAPAPGPTPDGRIYAARPGASHGDFPLTPHASPDRRVV-----SAPLGPLDPA 811
872 ENLDSLP-PKVPQREASL-----GPPGASLSQT 898
812 SAADGLPRFSPPTGSLRRRLGPHAPPAATLRT 846

```

RESULT 15
US-10-140-922-544

Sequence 544; Application US/10140922
Publication No. US20030138889A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Beresini, Laureen

APPLICANT: DeForge, Laura

APPLICANT: Desnoyers, Luc

APPLICANT: Filzaro, Ellen

APPLICANT: Gao, Wei-Qiang

APPLICANT: Gettisen, Mary E.

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Sherwood, Steven

APPLICANT: Smith, Victoria

APPLICANT: Stewate, Timothy A.

APPLICANT: Tumas, Daniel

APPLICANT: Watanabe, Colin K

APPLICANT: Wood, William

APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

TITLE OF INVENTION: ACTS ENCODING THE SAME

FILE REFERENCE: P33081C179

CURRENT FILING DATE: 2002-05-07

NUMBER OF SEQ ID NOS: 550

SEQ ID NO 544

LENGTH: 888

TYPE: PRT

ORGANISM: Homo Sapien

US-10-140-922-544

Query Match 36.9%; Score 2009.5; DB 12; Length 888;
Best Local Similarity 45.2%; Pred. No. 7.4e-160; Indels 141; Gaps 20;
Matches 423; Conservative 130; Mismatches 241

```

5 ALLYFTLLHPAGAPPEDESEPISSHGNYTKQYVFGHKGRNTTOR--HRLDIOMIM 62
12 ALLLLLLLGGAGHGFPEEPPLSVAPRDYLNHYVFGSGGRLLTPAEGADDLNIQVLT 71
63 INNGTLYIAARDHIYTVIDTSHTEIYCSKLLTKSKROADVDTCRMKGKHECHNEFK 122
72 RYNRTFLFGDRDNLRYVELLEPPTSTELKRYKRLTWRSNPSDINVCRMKGKQEGECRNFK 131
123 VLLKNDLALFYCGTNAENPSCRNKMDLEPFGEFSGMARCPYDAKHANVALFADGKL 182
132 VLLRDESTLFCVGSNAENPVCANYSIDLQGVGNISGMARCPYDPKHANVALFSDGML 191
183 YSATYTDFLAIDAVIYRSLGSPFLRTVYKHSKWLKEPFYQAVDYGDIYFFFEIAYE 242
192 FTAIYTDFLAIDAVIYRSLGDRPTLRIVYKHSKWLKEPFYQAVDYGDIYFFFEIAYE 251
243 YNTMGKVYFPRVAQYCKNDMGSGORVLEKQWTSFLKARLNCSPGDSHFYENILQAVTV 302
252 FNYLEKVVSVARVCKNDVGSPVLEKQWTSFLKARLNCSPGDSHFYENILQAVTV 311
303 IINGRDVLAIFSTPYNSIPSSAVCAVMDLDAVFTGRFEQKSPDSTWTPVDEVRP 362
312 VSLGGRPVVLAIFSTPYNSIPSSAVCAVMDLDAVFTGRFEQKSPDSTWTPVDEVRP 371
363 KRPFGCCAGSSSLERATYNSNEPDDTLNFKTHPLMDEAVPSIFNRPMWLFRTMVRRLTK 422
372 RRPFGCCAPGM--QYNASSALPDDTLNFKTHPLMDEAVPSIFNRPMWLFRTMVRRLTK 429
423 IAVDTAGFYONHTVVFELGSEKGIILKFLAR--IGNSGFLNSLPLEMSVYNSKCSYD 480
430 VAVDVAGFWGNGTIVVFELGSEAGTVLKFVLRNASTSGTSGVLFEEBETRYRDPDRCGP 489

```

QY 481 GVED--KRIMGMOULDRASSSLVVAFTVCIVKPLGRCERHCKCKTCIASRDPYCGWIK 538
 Db 490 GGETGQRLLSLELDNASSGGLAAPPRCVAVPVARCOQYSGCMKNCIGSDPYCGNAPD 549
 QY 539 GGACSHLSPNRLTFEODIERGNTDGLDCHNSFVALNGHSSSLLPSTTSDSTAGEGYE 598
 Db 550 -GSCIFLSPGTAAFEQDVSGASTGLDC----- 578
 QY 599 SRGGMIDWKHLLDSDSTDPGLGAVSNHODKKGVIRESYKGDQVPTLLAIVILA 658
 Db 579 -----TGLRLASLSEDRAGLVSNLVTSSVAA 606
 QY 659 FVMGAVFSGITV-XVCDDRRKDVAVQKKEKE--LTSRRGSMSSVTKLSGLFGDTQSK 715
 Db 607 FVVGAVVSGFVGVGTVGLRERELA--RRKXKAILAHAGAVILSVRL-----GERRAQ 660
 QY 716 DP-----KPEAILTPIMHNGKLATPGNTAK-MLIKADQHHLDLTALPTPEST 761
 Db 661 GPGGRGGGGGAGVPPFALLAPLMQNG-----WAKATLLGGPHDLDSGLLPTPEQT 713
 QY 762 PTLQOKRKP-----SRGSRERENONLINACTMDMPMGSPVPTDLP---LRASPSH 811
 Db 714 P-LPQKRLPTPHPHALGPRAMDH-----GHPILLPASASSSLLLLAPAR 757
 QY 812 IPSVYVLPITQOQYCHEYVDQPKMSEVAQMALEDQATLEKTIKEHLSKSPNHGVNLV 871
 Db 758 APPOPPAFGEPTPDGRLYAARPGRASHGDFILTPHASPDRRRV-----SAPTCGLDPA 811
 QY 872 ENLDSLP-PXVPQREASL-----GPPQASLSQT 898
 Db 812 SAADGLPRPMSPPFTGSLRPLGPHAPPAATLRR 846

Search completed: October 23, 2003, 17:20:47
 Job time : 90 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 23, 2003, 17:09:07 ; Search time 57 Seconds

(without alignments)
4663.056 Million cell updates/sec

Title: US-09-856-681-2

Perfect score: 5450

Sequence: 1 MRSKALLLYFTLHFAAGAF.....PKKSFAPLSTSMKKNDACT 1030

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

SPTREMBL 23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriopl:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5450	100.0	1030	4	Q9H2E6
2	5431.5	99.7	1049	4	Q9P2H9
3	5024.5	92.2	1005	11	Q9BQ71
4	3677	67.5	699	4	Q9GSM4
5	2980.5	54.7	574	4	Q9GSM9
6	2963	54.4	587	11	Q9BUT0
7	2961	54.3	605	11	Q9BXZ7
8	2952	54.2	562	4	Q9G5Y4
9	2947	54.1	562	4	Q9NC49
10	2621.5	48.1	1073	4	Q9GFO4
11	2368	43.4	1073	4	Q9NRY4
12	2349	43.1	1017	4	Q9NRY5
13	2315.5	42.5	998	4	Q9NRY6
14	2314	42.5	1022	4	Q9NRY3
15	2314	42.5	1022	4	Q9P249
16	2213	40.6	418	4	Q9GUT9

17	1908.5	35.0	595	11	Q9BKX8	Q9BKX8 mus musculus
18	1876.5	34.4	687	4	Q9BXR8	Q9BXR8 homo sapien
19	1611	29.6	476	4	Q9NRY7	Q9NRY7 homo sapien
20	1526.5	28.0	963	11	Q9LY36	Q9LY36 mus musculus
21	1516.5	27.8	935	4	Q9GTF8	Q9GTF8 homo sapien
22	1386.5	25.4	367	4	Q9H4H9	Q9H4H9 homo sapien
23	1355.5	24.9	923	11	Q9R4U3	Q9R4U3 mus musculus
24	1320	24.2	920	11	Q9R4U4	Q9R4U4 ratu mus norv
25	856.5	15.7	770	5	Q44253	Q44253 drosophila
26	854.5	15.7	770	5	Q9V3M4	Q9V3M4 drosophila
27	847.5	15.6	774	13	Q8ITW9	Q8ITW9 xenopus lae
28	823.5	15.1	844	11	Q9BXU8	Q9BXU8 mus musculus
29	819.5	15.0	754	4	Q8TDV7	Q8TDV7 homo sapien
30	817.5	15.0	748	4	Q8TB71	Q8TB71 homo sapien
31	802.5	14.7	777	11	Q8BWF6	Q8BWF6 mus musculus
32	802.5	14.7	777	11	Q8BH34	Q8BH34 mus musculus
33	779.5	14.3	724	5	Q9V7G7	Q9V7G7 drosophila
34	770.5	14.1	1202	4	Q9P283	Q9P283 homo sapien
35	768	14.0	697	5	Q8MLE1	Q8MLE1 drosophila
36	763	14.0	635	4	Q9GXX0	Q9GXX0 homo sapien
37	749	13.7	756	13	Q9GCU9	Q9GCU9 gallus gall
38	748	13.7	775	11	Q9GX23	Q9GX23 mus musculus
39	736.5	13.5	799	11	Q8BYC1	Q8BYC1 mus musculus
40	694.5	12.7	616	5	Q9V7P8	Q9V7P8 drosophila
41	693	12.7	963	4	Q9COC4	Q9COC4 homo sapien
42	686	12.6	782	4	Q9NS98	Q9NS98 homo sapien
43	681	12.5	893	4	Q9CDB8	Q9CDB8 homo sapien
44	671	12.3	761	4	Q8WU49	Q8WU49 homo sapien
45	666.5	12.2	1093	5	Q9VTT0	Q9VTT0 drosophila

ALIGNMENTS

RESULT 1	Q9H2E6	PREDIMINARY;	PRT; 1030 AA.
ID	Q9H2E6		
AC	Q9H2E6;		
DT	01-MAR-2001 (TREMBLrel. 16, Created)		
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)		
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)		
DE	Semaphorin SEMA6A1.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=20564339; PubMed=10933894;		
RA	Klostermann A., Lutz B., Gertler F., Behl C.;		
RT	"The orthologous human and murine semaphorin 6A-1 proteins (SEMA6A-1/SEMA6A-1) bind to the Enabled/Vasodilator-stimulated phosphoprotein-1 like Protein (EVL) via a novel carboxyl-terminal zyxin-like domain.";		
RL	J. Biol. Chem. 275:39647-39653(2000).		
DR	EMBL; AF279656; AAG29378.1; ..		
DR	InterPro; IPR003659; Plexin-like.		
DR	InterPro; IPR001627; Sema.		
DR	Pfam; PF01403; Sema; 1.		
DR	SMART; SMO0423; PSI; 1.		
DR	SMART; SMO0630; Sema; 1.		
DR	SMART; 1030 AA; 114368 MW; A57B79C10AEC4B34 CRC64;		
DR	SMART; 1030 AA; 114368 MW; A57B79C10AEC4B34 CRC64;		

QY	1	MRSKALLLYFTLHFAAGAFPPDSFPISSHNNTKQVFFGHHKRGRTTGRHRIDIQ 60
DB	1	MRSKALLLYFTLHFAAGAFPPDSFPISSHNNTKQVFFGHHKRGRTTGRHRIDIQ 60
QY	61	IMWNGTLYIARBDIYVDIDTSTTEIYSGKLTWRSRQADVDTCRMKGKHKECHNF 120
DB	61	IMWNGTLYIARBDIYVDIDTSTTEIYSGKLTWRSRQADVDTCRMKGKHKECHNF 120

```

QY 121 IKVLLKKNDALFVCGTNAFNPSCRYKMDTLEPFDEDESGMARCEYDAKHAVALFADG 180
Db 121 IKVLLKKNDALFVCGTNAFNPSCRYKMDTLEPFDEDESGMARCEYDAKHAVALFADG 180
QY 181 KLYSAVTDFLAIDAVIYISLGSPLRTYKHSKMLKEPYFOADVGYIYFFPREA 240
Db 181 KLYSAVTDFLAIDAVIYISLGSPLRTYKHSKMLKEPYFOADVGYIYFFPREA 240
QY 241 VEYNTMGKVPFPAVAVOCNDMGSGQVLEKQWTSFLKARLNSVGDSHFYENIIQAVT 300
Db 241 VEYNTMGKVPFPAVAVOCNDMGSGQVLEKQWTSFLKARLNSVGDSHFYENIIQAVT 300
QY 241 VEYNTMGKVPFPAVAVOCNDMGSGQVLEKQWTSFLKARLNSVGDSHFYENIIQAVT 300
Db 241 VEYNTMGKVPFPAVAVOCNDMGSGQVLEKQWTSFLKARLNSVGDSHFYENIIQAVT 300
QY 301 DVIRINGRDVVLATFSTPYNSIPGSAVCAVMDLJASVFTGRPEKQSPDSTWTPVDER 360
Db 301 DVIRINGRDVVLATFSTPYNSIPGSAVCAVMDLJASVFTGRPEKQSPDSTWTPVDER 360
QY 361 VPKPRPGCCAGSSSLERATSNFPDDTINFIKTHPLMDAVPSINRPFELTMYRYL 420
Db 361 VPKPRPGCCAGSSSLERATSNFPDDTINFIKTHPLMDAVPSINRPFELTMYRYL 420
QY 421 TKIAVDTAAGPYONHTVFLGSEKGIILKFLARIGNSGLNDLPLEMSVYNSKCSYD 480
Db 421 TKIAVDTAAGPYONHTVFLGSEKGIILKFLARIGNSGLNDLPLEMSVYNSKCSYD 480
QY 481 GVEDKRIWGQDLPRASSLYVAFSTCVIKVPLGRCEBHGCKCTCIASRDPYGMKEGG 540
Db 481 GVEDKRIWGQDLPRASSLYVAFSTCVIKVPLGRCEBHGCKCTCIASRDPYGMKEGG 540
QY 541 ACSHLSPNSRLTFEODIERGNTDGLGCHNSFVALNHSLSLPSSTTSSTQOEGYER 600
Db 541 ACSHLSPNSRLTFEODIERGNTDGLGCHNSFVALNHSLSLPSSTTSSTQOEGYER 600
QY 601 GGMLDWGHLDSPDSTPLGAVSHNHQDKGVIRESYLKGHDQVPLTLLAVALIAFV 660
Db 601 GGMLDWGHLDSPDSTPLGAVSHNHQDKGVIRESYLKGHDQVPLTLLAVALIAFV 660
QY 661 MGAVFSGITVYCVCDHRKQVAVVQKEKELTHSRGSMGSVTKLSGLFDDTQSKPKPE 720
Db 661 MGAVFSGITVYCVCDHRKQVAVVQKEKELTHSRGSMGSVTKLSGLFDDTQSKPKPE 720
QY 721 AILTEPLMHNGKLAIPGNTAKMLIKADQHLDTLALPTPESTPTLQOKRPSRGREREN 780
Db 721 AILTEPLMHNGKLAIPGNTAKMLIKADQHLDTLALPTPESTPTLQOKRPSRGREREN 780
QY 781 QNLINACTCKMPWGSVPIPTDPLRASPISHSVVLPITQOQYQEEYVDQRMSEVAQ 840
Db 781 QNLINACTCKMPWGSVPIPTDPLRASPISHSVVLPITQOQYQEEYVDQRMSEVAQ 840
QY 841 MALEDQAATLEYKTIKELHLSKSPNHGVNLVENIDSLPRKVPQREASLGPASLSQTL 900
Db 841 MALEDQAATLEYKTIKELHLSKSPNHGVNLVENIDSLPRKVPQREASLGPASLSQTL 900
QY 901 SKLEEMHSSSYGVYKRSYPTNSLTRSHQATTLKRNNTNSSNSHLSRNOFGRGNP 960
Db 901 SKLEEMHSSSYGVYKRSYPTNSLTRSHQATTLKRNNTNSSNSHLSRNOFGRGNP 960
QY 961 PAPORVDSICVHSSQPGGAATVSRQSLNAYNSLTRSGLKRTPSLKPDVPPKSPAPLS 1020
Db 961 PAPORVDSICVHSSQPGGAATVSRQSLNAYNSLTRSGLKRTPSLKPDVPPKSPAPLS 1020
QY 1021 TSMKPNDACT 1030
Db 1021 TSMKPNDACT 1030

```

```

DE Hypothetical protein KIAA1368 (Fragment).
GN KIAA1368.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OK NCBI_TaxId=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20181126; PubMed=10718198;
RA Nagase T., Kikuno R., Ishikawa K., Hirose M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XVI.
RT The complete sequences of 150 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Ref. 7:65-73(2000).
DR EMBL; AB037789; BAA92606.1; -.
DR Genbank; U003659; Pleixin-like.
DR InterPro; IPR001627; Pleixin-like.
DR Pfam; PF01403; Sema; 1.
DR SMART; SM00423; Sema; 1.
DR SMART; SM00630; Sema; 1.
DR Hypothetical protein.
FT NON TER.
SQ
SEQUENCE 1049 AA; 11651 MW; 7781D20AC07A8AEA CRC64;
Query Match 99.7%; Score 5431.5; DB 4; Length 1049;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 1030; Conservative 0; Mismatches 0; Indels 17; Gaps 1;
QY 1 MRSEALLFTLLHFAAGPEDESEPISTSHGNYTKQYFVYGHKGRNTQHRIDIQ 60
Db 3 MRSEALLFTLLHFAAGPEDESEPISTSHGNYTKQYFVYGHKGRNTQHRIDIQ 62
QY 61 IMINGTLYIARDDIYTVDDITSHTEIYCSKLTWKSROADVDTCKMKGKDECHNF 120
Db 63 IMINGTLYIARDDIYTVDDITSHTEIYCSKLTWKSROADVDTCKMKGKDECHNF 122
QY 121 IKVLLKKNDALFVCGTNAFNPSCRYKMDTLEPFDEDESGMARCEYDAKHAVALFADG 180
Db 121 IKVLLKKNDALFVCGTNAFNPSCRYKMDTLEPFDEDESGMARCEYDAKHAVALFADG 182
QY 181 KLYSAVTDFLAIDAVIYISLGSPLRTYKHSKMLKEPYFOADVGYIYFFPREA 240
Db 181 KLYSAVTDFLAIDAVIYISLGSPLRTYKHSKMLKEPYFOADVGYIYFFPREA 242
QY 241 VEYNTMGKVPFPAVAVOCNDMGSGQVLEKQWTSFLKARLNSVGDSHFYENIIQAVT 300
Db 241 VEYNTMGKVPFPAVAVOCNDMGSGQVLEKQWTSFLKARLNSVGDSHFYENIIQAVT 302
QY 301 DVIRINGRDVVLATFSTPYNSIPGSAVCAVMDLJASVFTGRPEKQSPDSTWTPVDER 360
Db 301 DVIRINGRDVVLATFSTPYNSIPGSAVCAVMDLJASVFTGRPEKQSPDSTWTPVDER 362
QY 361 VPKPRPGCCAGSSSLERATSNFPDDTINFIKTHPLMDAVPSINRPFELTMYRYL 420
Db 361 VPKPRPGCCAGSSSLERATSNFPDDTINFIKTHPLMDAVPSINRPFELTMYRYL 422
QY 421 TKIAVDTAAGPYONHTVFLGSEKGIILKFLARIGNSGLNDLPLEMSVYNSKCSYD 480
Db 421 TKIAVDTAAGPYONHTVFLGSEKGIILKFLARIGNSGLNDLPLEMSVYNSKCSYD 482
QY 481 GVEDKRIWGQDLPRASSLYVAFSTCVIKVPLGRCEBHGCKCTCIASRDPYGMKEGG 540
Db 481 GVEDKRIWGQDLPRASSLYVAFSTCVIKVPLGRCEBHGCKCTCIASRDPYGMKEGG 542
QY 541 ACSHLSPNSRLTFEODIERGNTDGLGCHNSFVALN-----GHSSSL 583
Db 541 ACSHLSPNSRLTFEODIERGNTDGLGCHNSFVALNISTPLPDNEMSYTVYGHSSSL 602
QY 584 PSTTSDSTQOEGYERSGMLDWKHLIDSDSTDPGLGAVSHNHQDKGVIRESYLKGHD 643
Db 603 PSTTSDSTQOEGYERSGMLDWKHLIDSDSTDPGLGAVSHNHQDKGVIRESYLKGHD 662

```

QY 644 QLVPTLLAIAVILAFVWGAVFSGITVYCVCDHRKRDVAVYQKKEKELTHSRGSMSSVT 703
 Db 663 QLVPTLLAIAVILAFVWGAVFSGITVYCVCDHRKRDVAVYQKKEKELTHSRGSMSSVT 722
 QY 704 KLSGLFGDTQSDKDPREALILPLMNGKLAIPGNTAKMLIKADQHLLDLTALPTBESTPT 763
 Db 723 KLSGLFGDTQSDKDPREALILPLMNGKLAIPGNTAKMLIKADQHLLDLTALPTBESTPT 782
 QY 764 LQCKKPSRSGRSEWRNQNLIINACTKMPMGSPIVITDPLPRASPSHIPVVLVLPITQQ 823
 Db 783 LQCKKPSRSGRSEWRNQNLIINACTKMPMGSPIVITDPLPRASPSHIPVVLVLPITQQ 842
 QY 824 GYQHEVYVQPKKSEVAQWALBDQATLEKTIKEHLSKSPNHGVLNVLNDSLPKXPQ 883
 Db 843 GYQHEVYVQPKKSEVAQWALBDQATLEKTIKEHLSKSPNHGVLNVLNDSLPKXPQ 902
 QY 884 REASLGPPGASISQGTLSKRLMHHSSSYGVYKRSYPTNSLTRSHQATTLKRNNTSSN 943
 Db 903 REASLGPPGASISQGTLSKRLMHHSSSYGVYKRSYPTNSLTRSHQATTLKRNNTSSN 962
 QY 944 SSHLSRNSQFGRGNDPPAPQVDSIQVHSSQPSGOAVTSRQPSLAINVNSLTRSGLKRT 1003
 Db 963 SSHLSRNSQFGRGNDPPAPQVDSIQVHSSQPSGOAVTSRQPSLAINVNSLTRSGLKRT 1022
 QY 1004 PSLKPDVPPKPSFAPLSTSMKENDACT 1030
 Db 1023 PSLKPDVPPKPSFAPLSTSMKENDACT 1049

RESULT 3

Q9E071 PRELIMINARY; PRT; 1005 AA.

AC Q9E071: 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Axon guidance signal SEMA6A.
 GN SEMA6A.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_Taxid:10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA MEDLINE=20364339; PubMed=10993894;
 RX Klosternann A., Lutz B., Gertler F., Behl C.;
 RT "The orthologous human and murine semaphorin 6A-1 proteins (SEMA6A-1/Sema6A-1) bind to the disabled/Vasodilator-stimulated Phosphoprotein-1-like protein (EVL) via a novel carboxyl-terminal zyxin-like domain.";
 RT J. Biol. Chem. 275:39647-39653 (2000).
 RL EMBL; AF288666; AAC29494.1; -;
 DR EMBL; AF288666; AAC29494.1; -;
 DR MGD; MGI:1203727; Sema6a.
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00423; PST; 1.
 DR SMART; SM00630; Sema; 1.
 SQ SEQUENCE 1005 AA; 111758 MW; 57B69927F45B079D CRC64;

Query Match 92.2%; Score 5024.5; DB 11; Length 1005;

Best Local Similarity 91.9%; Pred. No. 0;
 Matches 948; Conservative 26; Mismatches 30; Indels 27; Gaps 2;

QY 1 MRSEALLVYTLIFACAGFPEDESEPTISGHANYTQYPIVFGHKGRNTTQHRDLIDQ 60
 Db 1 MRPAALLCTLLHAGAGFPEDESEPTISGHANYTQYPIVFGHKGRNTTQHRDLIDQ 60
 QY 61 IMYNGVLYAARDHITYVDITDSSHTEIYCSKKITMKSGQAVDCCRMKGKHKDCQNF 120
 Db 61 IMYNGVLYAARDHITYVDITDSSHTEIYCSKKITMKSGQAVDCCRMKGKHKDCQNF 120
 Db 61 IMYNGVLYAARDHITYVDITDSSHTEIYCSKKITMKSGQAVDCCRMKGKHKDCQNF 120

QY 121 IKVLLKNDALFVCGTNAFNPSCRNRYKMDTLPEFGDEFSGMARCPYDAKHANVALPADG 180
 Db 121 IKVLLKNDALFVCGTNAFNPSCRNRYKMDTLPEFGDEFSGMARCPYDAKHANVALPADG 180
 QY 181 KLYGATYDPLAIDAVIYRSLGSEPTLRTYVHDSKWLKEPFYQAVDYGDIYFFFRRIA 240
 Db 181 KLYGATYDPLAIDAVIYRSLGSEPTLRTYVHDSKWLKEPFYQAVDYGDIYFFFRRIA 240
 QY 241 VEYNTMGKVPFPRVAQVCKNDMGSGQVLEKQWTSFLKARLNCVPGDSHFYFNILQAVT 300
 Db 241 VEYNTMGKVPFPRVAQVCKNDMGSGQVLEKQWTSFLKARLNCVPGDSHFYFNILQAVT 300
 QY 301 DVIRINGRDVTLAFSTPYNSIPGSAVAYMDLISFTGRFQKSPDSTWTVPDER 360
 Db 301 DVIRINGRDVTLAFSTPYNSIPGSAVAYMDLISFTGRFQKSPDSTWTVPDER 360
 QY 361 VPKRPGCCAGSSSLERATSNPEPDDPLNFKTHPLMDEAVPSIFNSPMFLRTWRYRL 420
 Db 361 VPKRPGCCAGSSSLERATSNPEPDDPLNFKTHPLMDEAVPSIFNSPMFLRTWRYRL 420
 QY 421 TKIADTAAGFYQNHVYFLGSEKGIILKFLARIGNSGFLNDSLFLEEMSVYNSKCYD 480
 Db 421 TKIADTAAGFYQNHVYFLGSEKGIILKFLARIGNSGFLNDSLFLEEMSVYNSKCYD 480
 QY 481 GVEDKRIHQMDLRASSSLYYAFSTCVYKVLGRCHRGKCKTCIASRDPGCMIXEG 540
 Db 481 GVEDKRIHQMDLRASSSLYYAFSTCVYKVLGRCHRGKCKTCIASRDPGCMIXEG 540
 QY 541 ACSHLSFNSRLTFBODIERGNTDGLGDCNHFVALNGHSSSLPSTTSDSIAOGEYSR 600
 Db 541 ACSHLSFNSRLTFBODIERGNTDGLGDCNHFVALNGHSSSLPSTTSDSIAOGEYSR 600
 QY 575 SCALHSPSKLTFBODIERGNTDGLGDCNHFVALNGHSSSLPSTTSDSIAOGEYSR 634
 Db 575 SCALHSPSKLTFBODIERGNTDGLGDCNHFVALNGHSSSLPSTTSDSIAOGEYSR 634
 QY 601 GMDLWKMILDSPOSTDLGAVSSHNDKGVIRESYLKAGHDQVPTLLAIAVILAFV 660
 Db 601 GMDLWKMILDSPOSTDLGAVSSHNDKGVIRESYLKAGHDQVPTLLAIAVILAFV 660
 QY 661 MGAVVSGITVYCVCDHRKRDVAVYQKKEKELTHSRGSMSSVTXLSGFGDTQSDPKPE 720
 Db 661 MGAVVSGITVYCVCDHRKRDVAVYQKKEKELTHSRGSMSSVTXLSGFGDTQSDPKPE 720
 QY 694 MGAVVSGITVYCVCDHRKRDVAVYQKKEKELTHSRGSMSSVTXLSGFGDTQSDPKPE 720
 Db 694 MGAVVSGITVYCVCDHRKRDVAVYQKKEKELTHSRGSMSSVTXLSGFGDTQSDPKPE 720
 QY 721 AILPLMNGKLAIPGNTAKMLIKADQHLLDLTALPTBESTPTLQCKRPSRSGRSEWRN 780
 Db 721 AILPLMNGKLAIPGNTAKMLIKADQHLLDLTALPTBESTPTLQCKRPSRSGRSEWRN 780
 QY 754 AILPLMNGKLAIPGNTAKMLIKADQHLLDLTALPTBESTPTLQCKRPSRSGRSEWRN 780
 Db 754 AILPLMNGKLAIPGNTAKMLIKADQHLLDLTALPTBESTPTLQCKRPSRSGRSEWRN 780
 QY 781 QNLIINACTKMPMGSPIVITDPLPRASPSHIPVVLVLPITQQGYQHEVYVDPKXSEVA 839
 Db 781 QNLIINACTKMPMGSPIVITDPLPRASPSHIPVVLVLPITQQGYQHEVYVDPKXSEVA 839
 QY 840 QNALDQATLEKTIKEHLSKSPNHGVLNVLNDSLPKXPORBASLGPPGASISQGT 899
 Db 840 QNALDQATLEKTIKEHLSKSPNHGVLNVLNDSLPKXPORBASLGPPGASISQGT 899
 QY 874 QNALDQATLEKTIKEHLSKSPNHGVLNVLNDSLPKXPORBASLGPPGASISQGT 899
 Db 874 QNALDQATLEKTIKEHLSKSPNHGVLNVLNDSLPKXPORBASLGPPGASISQGT 899
 QY 900 LSKRLMHHSSSYGVYKRSYPTNSLTRSHQATTLKRNNTSSNSHLSRNSQFGRGNDP 959
 Db 900 LSKRLMHHSSSYGVYKRSYPTNSLTRSHQATTLKRNNTSSNSHLSRNSQFGRGNDP 959
 QY 934 LSKRLMHHSSSYGVYKRSYPTNSLTRSHQATTLKRNNTSSNSHLSRNSQFGRGNDP 959
 Db 934 LSKRLMHHSSSYGVYKRSYPTNSLTRSHQATTLKRNNTSSNSHLSRNSQFGRGNDP 959
 QY 960 PPAPQVDSIQVHSSQPSGOAVTVSRQPSLAINVNSLTRSGLKRTSLKPDVPPKPSFAPL 1019
 Db 960 PPAPQVDSIQVHSSQPSGOAVTVSRQPSLAINVNSLTRSGLKRTSLKPDVPPKPSFAPL 1019
 QY 994 PPAPQVDSIQVHSSQPSGOAVTVSRQPSLAINVNSLTRSGLKRTSLKPDVPPKPSFAPL 1019
 Db 994 PPAPQVDSIQVHSSQPSGOAVTVSRQPSLAINVNSLTRSGLKRTSLKPDVPPKPSFAPL 1019
 QY 1020 STSKMPNDACT 1030
 Db 1020 STSKMPNDACT 1030
 QY 995 STSKMPNDACT 1005
 Db 995 STSKMPNDACT 1005

RESULT 4

Q96SW4 PRELIMINARY; PRT; 699 AA.

ID Q96SW4: 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Hypothetical protein Flj14595.

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominda; Homo.
 ON NCBI TaxID=9606;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
 RA Magatsuna M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saico K.,
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuno Y.,
 RA Niimiya K., Iwayanagi T.,
 RT "NEO human cDNA sequencing project."
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK027501; BAB55158.1;
 DR InterPro: IPR003659; Plectin-like.
 DR InterPro: IPR001627; Sema.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00423; PSI; 1.
 KM Hypothetical protein.
 SQ SEQUENCE 699 AA; 76723 MW; 2E5F11D59741394 CRC64;
 Query Match 67.5%; Score 3677; DB 4; Length 699;
 Best Local Similarity 100.0%; Pred. No. 1.5e-290; Indels 0; Gaps 0;
 Matches 699; Conservative 0; Mismatches 0;
 QY 332 MLDIASVFTGRFKEOKSPDSTWTPVPPDERVPPKPPGCGAGSSSLERVAISNEFPDPTLNF 391
 DB 1 MLDIASVFTGRFKEOKSPDSTWTPVPPDERVPPKPPGCGAGSSSLERVAISNEFPDPTLNF 60
 QY 392 ICHPLMDEAVPSIFNRPWFLRTWRYRLTKIADVTAGPYQNTVPLSGEKKILKFL 451
 DB 61 ICHPLMDEAVPSIFNRPWFLRTWRYRLTKIADVTAGPYQNTVPLSGEKKILKFL 120
 QY 452 ARGNSEFLMDSLFLEMSVYNSKSGYDGEDKRIKMGQLDRASSSLVYAFSTCVIKVP 511
 DB 121 ARGNSEFLMDSLFLEMSVYNSKSGYDGEDKRIKMGQLDRASSSLVYAFSTCVIKVP 180
 QY 512 LGRCEHKGCKKCTCIASRDPCYCGWIKGAGCSHLSFNSRLTFEODIERGNTDGLGDCNS 571
 DB 181 LGRCEHKGCKKCTCIASRDPCYCGWIKGAGCSHLSFNSRLTFEODIERGNTDGLGDCNS 240
 QY 572 FVALNGSSSLPSTTSDSTAOEGYSGRGMLDWKHLSDPSDPLGAVSSHNHODKK 631
 DB 241 FVALNGSSSLPSTTSDSTAOEGYSGRGMLDWKHLSDPSDPLGAVSSHNHODKK 300
 QY 632 GVIESYIKGHQDLVPTLLAIVLAFVWGA VFSGITVYCVCDHRKDVAVVQREKEL 691
 DB 301 GVIESYIKGHQDLVPTLLAIVLAFVWGA VFSGITVYCVCDHRKDVAVVQREKEL 360
 QY 692 THSRGSSSVTKLSLFGDTOSKDPKPEAILPLMHNGKLTATGNTAKMLIKADQHLD 751
 DB 361 THSRGSSSVTKLSLFGDTOSKDPKPEAILPLMHNGKLTATGNTAKMLIKADQHLD 420
 QY 752 LTAIPFESPTTLOOKRPSRSGREWRNOMLINA CTKMPGSGVPIPTDPLRASPSH 811
 DB 421 LTAIPFESPTTLOOKRPSRSGREWRNOMLINA CTKMPGSGVPIPTDPLRASPSH 480
 QY 812 IPSVYVLPITQOGYQHEHYVDQPKMSVYVQALMLEDQAATLEKTIKEHLSKSPNHGVLV 871
 DB 481 IPSVYVLPITQOGYQHEHYVDQPKMSVYVQALMLEDQAATLEKTIKEHLSKSPNHGVLV 540
 QY 872 ENLDSLPRKVPQREASIGPPGASISQGTGLSKRLEMHSSSYGVYKSYPTNSLTRSHQA 931
 DB 541 ENLDSLPRKVPQREASIGPPGASISQGTGLSKRLEMHSSSYGVYKSYPTNSLTRSHQA 600
 QY 932 TTLKRNNTSSNSHLSRNSQSPRGNDNPPAPORVDSIQVHSSQSPGQAVTVSRQPSLNA 991
 DB 601 TTLKRNNTSSNSHLSRNSQSPRGNDNPPAPORVDSIQVHSSQSPGQAVTVSRQPSLNA 660
 QY 992 YNSLTRSGIKRTPSLKPDVPPKPSFAPLSTSMKPNDACT 1030

DB 661 YNSLTRSGIKRTPSLKPDVPPKPSFAPLSTSMKPNDACT 699
 RESULT 5
 Q96SM8 PRELIMINARY; PRT; 574 AA.
 AC Q96SM8;
 DT 01-DEC-2001 (TRENBLREL. 19, Created)
 DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)
 DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)
 DE Hypothetical protein F141478.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominda; Homo.
 ON NCBI TaxID=9606;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
 RA Magatsuna M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saico K.,
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuno Y.,
 RA Niimiya K., Iwayanagi T.,
 RT "NEO human cDNA sequencing project."
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK027501; BAB55269.1;
 DR InterPro: IPR003659; Plectin-like.
 DR SMART; SM00423; PSI; 1.
 KM Hypothetical protein.
 SQ SEQUENCE 574 AA; 62822 MW; 0C79E01A4117A495 CRC64;
 Query Match 54.7%; Score 2980.5; DB 4; Length 574;
 Best Local Similarity 90.7%; Pred. No. 6.6e-234; Indels 59; Gaps 1;
 Matches 574; Conservative 0; Mismatches 0;
 QY 398 MDEAVPSIFNRPWFLRTWRYRLTKIADVTAGPYQNTVPLSGEKKILKFLARIGNS 457
 DB 1 MDEAVPSIFNRPWFLRTWRYRLTKIADVTAGPYQNTVPLSGEKKILKFLARIGNS 20
 QY 458 GFINDSLFLEMSVYNSKSGYDGEDKRIKMGQLDRASSSLVYAFSTCVIKVPLGCR 517
 DB 21 -----CSYDGEDKRIKMGQLDRASSSLVYAFSTCVIKVPLGCR 61
 QY 518 HGCKKCTCIASRDPCYCGWIKGAGCSHLSFNSRLTFEODIERGNTDGLGDCNSFYALNG 577
 DB 62 HGCKKCTCIASRDPCYCGWIKGAGCSHLSFNSRLTFEODIERGNTDGLGDCNSFYALNG 121
 QY 578 HSSSLPSTTSDSTAOEGYSGRGMLDWKHLSDPSDPLGAVSSHNHODKKVITRES 637
 DB 122 HSSSLPSTTSDSTAOEGYSGRGMLDWKHLSDPSDPLGAVSSHNHODKKVITRES 181
 QY 638 YLKGHQLVPTLLAIVLAFVWGA VFSGITVYCVCDHRKDVAVVQREKELTHSRG 697
 DB 182 YLKGHQLVPTLLAIVLAFVWGA VFSGITVYCVCDHRKDVAVVQREKELTHSRG 241
 QY 698 THSRGSSSVTKLSLFGDTOSKDPKPEAILPLMHNGKLTATGNTAKMLIKADQHLDLPT 757
 DB 242 THSRGSSSVTKLSLFGDTOSKDPKPEAILPLMHNGKLTATGNTAKMLIKADQHLDLPT 301
 QY 758 PESTPTLQOKRPSRSGREWRNOMLINA CTKMPGSGVPIPTDPLRASPSH 817
 DB 302 PESTPTLQOKRPSRSGREWRNOMLINA CTKMPGSGVPIPTDPLRASPSH 361
 QY 818 LPTQOGYQHEHYVDQPKMSVYVQALMLEDQAATLEKTIKEHLSKSPNHGVLVENDSL 877
 DB 362 LPTQOGYQHEHYVDQPKMSVYVQALMLEDQAATLEKTIKEHLSKSPNHGVLVENDSL 421
 QY 878 PPKVPQREASIGPPGASISQGTGLSKRLEMHSSSYGVYKSYPTNSLTRSHQATTLKRN 937
 DB 422 PPKVPQREASIGPPGASISQGTGLSKRLEMHSSSYGVYKSYPTNSLTRSHQATTLKRN 481
 QY 938 NTSNSNSHLSRNSQSPRGNDNPPAPORVDSIQVHSSQSPGQAVTVSRQPSLNA 997

```

Db      482 NTNSNSNSHSLSRNOSFGRGDPPPPAPORVDSIQVHSSOPSPSQAVTVSRQSLNAINSLTR 541
Qy      998 SGLKRTPSLKEDVPKPSFAPLSTSMKENDACT 1030
Db      542 SGLKRTPELKPDPVKPSFAPLSTSMKENDACT 574

RESULT 6
Q8BUTO
ID      Q8BUTO      PRELIMINARY;      PRT;      587 AA.
AC      Q8BUTO:
DT      01-MAR-2003 (TReMBLrel. 23, Created)
DT      01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT      01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE      Sema domain.
OS      Mus musculus (Mouse) .
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX      NCBI_TaxId=10090;
      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6J; TISSUE=Cerebellum;
RC      MEDLINE=22354683; PubMed=12466851;
RA      The FANTOM Consortium,
RA      the RIKEN Genome Exploration Research Group Phase I & II Team;
RT      "Analysis of the mouse transcriptome based on functional annotation of
      60,770 full-length cDNAs.";
RL      Nature 420:563-573(2002).
DR      EMBU; AK082711; BAC3682.1; --
SQ      SEQUENCE 587 AA; 66044 MW; E99C7B102C4DB97A CRC64;

```

Query Match	54.4%	Score 2963	DB 11	Length 587
Best Local Similarity	95.1%	Pred. No. 1.8e-232		
Matches 549	Conservative 13	Mismatches 15	Indels 0	Gaps 0
QY	1	MSEALLLYFTLLHAGAGFEPDESEPISSHGNYTKQYFVFGHKGGRNTYQRHLDIOM	60	
Db	1	MRPALLLICTLLHCGAGFEPDESEPISSHGNYTKQYFVFGHKGGRNTYQRHLDIOM	60	
QY	61	IMINNGTLYIAARDHIYYTVDITDTSHTBEIYCKKLTWKSROADVOTCRKGGKHQECNF	120	
Db	61	IMINRTLYVAARDHIYYTVDITDTSHTBEIYCKKLTWKSROADVOTCRKGGKHQECNF	120	
QY	121	IKVLLKNDDAFVCGTNAFNSPCNNYKMDLEPFGEDESGARCPYAKAHNVALLFADG	180	
Db	121	IKVLLKNDDDTLFVCGTNAFNSPCNNYRDTLETGDEDESGARCPYAKAHNVALLFADG	180	
QY	181	KLVSATVTDFLAIDAVIYRSLGJESPTLRTVKHSKMLKEPYFVQAVDGYDYIFFPREIA	240	
Db	181	KLVSATVTDFLAIDAVIYRSLGDSPTLRTVKHSKMLKEPYFVQAVDGYDYIFFPREIA	240	
QY	241	VEYNIMGKVFFPRVAOVCKNDMGSGRVLEKOWTSLTKRLNCSVPGDSHFENLLQAT	300	
Db	241	VEYNIMGKVFFPRVAOVCKNDMGSGRVLEKOWTSLTKRLNCSVPGDSHFENLLQAT	300	
QY	301	DVIRINGRDVVALATSTFPYNSIPGSVACAYMDLIVATGTFRKEQSDSTMTVPDDR	360	
Db	301	DVIRINGRDVVALATSTFPYNSIPGSVACAYMDLIVATGTFRKEQSDSTMTVPDDR	360	
QY	361	VPKPRECCAGSSSLERYATSNFEPDDTLNFIKTHPLMDEAVPSIFNRPMFLRTMYRYL	420	
Db	361	VPKPRECCAGSSSLERYATSNFEPDDTLNFIKTHPLMDEAVPSIFNRPMFLRTMYRYL	420	
QY	421	TKIADVTLAAGPQNHVVVFLGSKGILLKFLARIGNSGLNLSFLEEMSVNSEKSYD	480	
Db	421	TKIADVNAAGPQNHVVVFLGSEKGIILLKFLAIGSGFLNLSFLEEMVNVNPEKCSYD	480	
QY	481	GVEDKIMQMOIDRASSLYVAFSTCVIKVPLRCRCHSGCKKCTCLASDPYCGWIKBG	540	
Db	481	GVEDKIMQMOIDRAGSLYVAFSTCVIKVPLRCRCHSGCKKCTCLASDPYCGWIREG	540	
QY	541	ACGSHSPNSRLTEODIERGNTDGLDCCHNSFYALNG	577	

```

Db          541 SCALHSP.LSR.ITFEQDIERGNTDGLSDCHNS.FYALNG 577
RESULT 7
QBXXZ7      PRELIMINARY;      PRT;      605 AA.
QBXXZ7
AC   QBXXZ7.
DT   01-MAR-2003 (TREMBlrel. 23. Created)
DT   01-MAR-2003 (TREMBlrel. 23. Last sequence update)
DJ   01-MAR-2003 (TREMBlrel. 23. Last annotation update)
DE   Sema domain.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxId=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK042751; BAC1351.1; -.
SQ   SEQUENCE      605 AA;  68299 MW;  BAFE713BE8AFE90 CRC64;
Query Match      54.3%; Score 2961; DB 11; Length 605;
Best Local Similarity    94.5%; Pred. No.2.8e-23;
Matches 550; Conservative 14; Mismatches 10; Indels 0; Gaps 0
```

QY	I	NRSEALLIYFTLLHRAAGFPEDSPISISHGNATKYQVPAVGHKRGNTTORHLLDIOM	60
Db	1	MRFAALLIOLTLHOCAGAFPEDESPISISHGNATKYQVPAVGHKRGNTTORHLLDIOM	60
QY	61	IMINMGTLIYAARHIIYVDIDTSHTEETIYCSKLTWMSROADVDTCKMGKHKECHNF	120
Db	61	IMINMGTLIYAARHIIYVDIDTSHTEETIYCSKLTWMSROADVDTCKMGKHKECHNF	120
QY	121	IKVLLKRNDDALFVCGTNAFNPSCBNYMDLTLEPGEFSGMARCEYDAKANVALFADG	180
Db	121	IKVLLKRNDDTLFVCGTNAFNPSCBNYMDLTLEPGEFSGMARCEYDAKANVALFADG	180
QY	181	KLYSATVDFLAIDAVIYRSLGSESTLTLYGDSWMLKEPYVQAVDGDYIYFFREJA	240
Db	181	KLYSATVDFLAIDAVIYRSLGSESTLTLYGDSWMLKEPYVQAVDGDYIYFFREJA	240
QY	241	VEYNMGKVPFRVAVOVCKNDMGSGORVLEKOWTISFLKARJLNCSPVGBSHFEYNLQAVT	300
Db	241	VEYNMGKVPFRVAVOVCKNDMGSGORVLEKOWTISFLKARJLNCSPVGBSHFEYNLQAVT	300
QY	301	DVIRINGRDVVUATESTPYNSIPGSAVCAYDMLDIASVFTGPFKEQKSPDSTWTVPPDER	360
Db	301	DVIRINGRDVVUATESTPYNSIPGSAVCAYDMLDIANVFTGPFKEQKSPDSTWTVPPDER	360
QY	361	VKPRPGCCAGSSSLEKATVNSNEFPDDTLNFKTHPPLMDEANPSIFENRWEJRTWRYXL	420
Db	361	VKPRPGCCAGSSSLEKATVNSNEFPDDTLNFKTHPPLMDEAPSIINRWEJRTWRYXL	420
QY	421	TKIAVDLTAGPQNHVVFVLGSEKGIILFKTLARJNSGFLNDSLPLEEMSVMNSEKCSYD	480
Db	421	TKIAVDNAGPQNHVVFVLGSEKGIILFKTLARJNSGFLNDSLPLEEMNVVNPEKCSYD	480
QY	481	GVEDKALIMQDLDRASSLYVAFSTCVIKVPLGRCERHKKCKTCLASDDPYCGWIKKEG	540
Db	481	GVEDKALIMQDLDRASSLYVAFSTCVIKVPLGRCERHKKCKTCLASDDPYCGWVRESG	540
QY	541	ACGSHSPNRKLTPEODIERGNTDGLGDCNNSVVALNGHSSSL	582
Db	541	ACGSHSPNRKLTPEODIERGNTDGLGDCNNSVVALNDSTPL	582

RESULT 8

ID Q96SY4 PRELIMINARY; PRT: 562 AA.

AC Q96SY4; 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Hypothetical protein FLN14555.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]

SEQUENCE FROM N.A.
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
 RA Watanabe M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
 RA Matanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
 RA Yamamoto J., Makematsu A., Nakamura Y., Nagahara K., Masuno Y.,
 RA Nimomiya K., Iwayanagi T.;
 RT "NEBO human cDNA sequencing project."
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK027471; BAB55136.1; -;
 DR InterPro: IPR003659; Plexin-like.
 DR SMART; SM00423; PSI; 1.
 KW Hypothetical protein.

SQ SEQUENCE 562 AA; 61313 MW; 6AB3685FAD1DD78A CRC64;

Query Match 54.2%; Score 2952; DB 4; Length 562;
 Best Local Similarity 100.0%; Pred. No. 1.3e-231;
 Matches 562; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 469 MSVNSEKSYDGVEDKRMGMQLDRASSSLVVAESTCYIKYPLGRCERHGCKKTCIAS 528
 DB 1 MSVNSEKSYDGVEDKRMGMQLDRASSSLVVAESTCYIKYPLGRCERHGCKKTCIAS 60
 QY 529 RDPYCGWIKKEGACSHLSPNSRLTFEODIERGNTDGLGDCNHSFVALNHSSSLPSTT 588
 DB 61 RDPYCGWIKKEGACSHLSPNSRLTFEODIERGNTDGLGDCNHSFVALNHSSSLPSTT 120
 QY 589 SDSTAQEGYSGRGMLDWKHLSDPSDTPDLGAVSSHNDKKGVIRESYLKGHDQLVYV 648
 DB 121 SDSTAQEGYSGRGMLDWKHLSDPSDTPDLGAVSSHNDKKGVIRESYLKGHDQLVYV 180
 QY 649 TLTAIAVILAFVWGAVFSGITVYCVCDHRKRVAVVORKEKELTSRRSGMSVTKLSGL 708
 DB 181 TLTAIAVILAFVWGAVFSGITVYCVCDHRKRVAVVORKEKELTSRRSGMSVTKLSGL 240
 QY 709 FGDTSQDKPREALITPLMHNGLATPGNTAKMLIKADQHLDLTALPPESTPTLQOKR 768
 DB 241 FGDTSQDKPREALITPLMHNGLATPGNTAKMLIKADQHLDLTALPPESTPTLQOKR 300
 QY 769 KPSRSGREWERNQNLINACTKMPMGSPVITPDLPLRASPSPHPSVVLPTTQOQYHE 828
 DB 301 KPSRSGREWERNQNLINACTKMPMGSPVITPDLPLRASPSPHPSVVLPTTQOQYHE 360
 QY 829 YVDQPKMSVQAQALDQAATLEKTKIKELSSKSPHGVNIVENDLSLPKYPQREASL 888
 DB 361 YVDQPKMSVQAQALDQAATLEKTKIKELSSKSPHGVNIVENDLSLPKYPQREASL 420
 QY 889 GPPGASLSQGLSKRLEMHSSSYGVYKRSYPTNSLTRSHQATTLKRNNTNSNSHLS 948
 DB 421 GPPGASLSQGLSKRLEMHSSSYGVYKRSYPTNSLTRSHQATTLKRNNTNSNSHLS 480
 QY 949 RNQSPFGKGNPPAPQPVDSIQVHSSQPSGQAVTVSRQPSLAINVSLTRSGLKRTPSLKP 1008
 DB 481 RNQSPFGKGNPPAPQPVDSIQVHSSQPSGQAVTVSRQPSLAINVSLTRSGLKRTPSLKP 540
 QY 1009 DVPPKPSFAPLSTSMKENDACT 1030
 DB 541 DVPPKPSFAPLSTSMKENDACT 562

RESULT 9

ID Q8NC49 PRELIMINARY; PRT: 562 AA.

AC Q8NC49; 01-OCT-2002 (TREMBLrel. 22, Created)
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
 DE Hypothetical protein FLJ90494.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]

SEQUENCE FROM N.A.
 RA Isogai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
 RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,
 RA Yamamoto J., Makematsu A., Nakamura Y., Kojima S., Nagahara K.,
 RA Masuno Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
 RA Hattori A., Okumura K., Iwayanagi T., Nimomiya K.;
 RT "NEBO human cDNA sequencing project."
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK074975; BAC11326.1; -;
 DR Hypothetical protein.
 KW Hypothetical protein.

Query Match 54.1%; Score 2947; DB 4; Length 562;
 Best Local Similarity 99.8%; Pred. No. 3.4e-231;
 Matches 561; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 469 MSVNSEKSYDGVEDKRMGMQLDRASSSLVVAESTCYIKYPLGRCERHGCKKTCIAS 528
 DB 1 MSVNSEKSYDGVEDKRMGMQLDRASSSLVVAESTCYIKYPLGRCERHGCKKTCIAS 60
 QY 529 RDPYCGWIKKEGACSHLSPNSRLTFEODIERGNTDGLGDCNHSFVALNHSSSLPSTT 588
 DB 61 RDPYCGWIKKEGACSHLSPNSRLTFEODIERGNTDGLGDCNHSFVALNHSSSLPSTT 120
 QY 589 SDSTAQEGYSGRGMLDWKHLSDPSDTPDLGAVSSHNDKKGVIRESYLKGHDQLVYV 648
 DB 121 SDSTAQEGYSGRGMLDWKHLSDPSDTPDLGAVSSHNDKKGVIRESYLKGHDQLVYV 180
 QY 649 TLTAIAVILAFVWGAVFSGITVYCVCDHRKRVAVVORKEKELTSRRSGMSVTKLSGL 708
 DB 181 TLTAIAVILAFVWGAVFSGITVYCVCDHRKRVAVVORKEKELTSRRSGMSVTKLSGL 240
 QY 709 FGDTSQDKPREALITPLMHNGLATPGNTAKMLIKADQHLDLTALPPESTPTLQOKR 768
 DB 241 FGDTSQDKPREALITPLMHNGLATPGNTAKMLIKADQHLDLTALPPESTPTLQOKR 300
 QY 769 KPSRSGREWERNQNLINACTKMPMGSPVITPDLPLRASPSPHPSVVLPTTQOQYHE 828
 DB 301 KPSRSGREWERNQNLINACTKMPMGSPVITPDLPLRASPSPHPSVVLPTTQOQYHE 360
 QY 829 YVDQPKMSVQAQALDQAATLEKTKIKELSSKSPHGVNIVENDLSLPKYPQREASL 888
 DB 361 YVDQPKMSVQAQALDQAATLEKTKIKELSSKSPHGVNIVENDLSLPKYPQREASL 420
 QY 889 GPPGASLSQGLSKRLEMHSSSYGVYKRSYPTNSLTRSHQATTLKRNNTNSNSHLS 948
 DB 421 GPPGASLSQGLSKRLEMHSSSYGVYKRSYPTNSLTRSHQATTLKRNNTNSNSHLS 480
 QY 949 RNQSPFGKGNPPAPQPVDSIQVHSSQPSGQAVTVSRQPSLAINVSLTRSGLKRTPSLKP 1008
 DB 481 RNQSPFGKGNPPAPQPVDSIQVHSSQPSGQAVTVSRQPSLAINVSLTRSGLKRTPSLKP 540
 QY 1009 DVPPKPSFAPLSTSMKENDACT 1030
 DB 541 DVPPKPSFAPLSTSMKENDACT 562

RESULT 10

ID Q96T04

ID Q96T04 PRELIMINARY; PRT; 507 AA.
 AC Q96T04:
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein FL14533.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Isegai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
 RA Watanabe M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
 RA Yamamoto J., Nakamatsu A., Nakamura Y., Nagahara K., Masuno Y.,
 RA Nimomiyu K., Iwayanagi T.;
 RA "NEO human cDNA sequencing project";
 RT Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.
 RL EMBL, AK027439; BAB511.1;
 DR InterPro: IPR003659; Pfam: PF01403; SMART: SM00423; PSI: 1.
 DR SMART: SM00423; PSI: 1.
 KW Hypothetical protein.
 SQ SEQUENCE 507 AA; 55464 MW; 8CC567B43BC51B3 CRC64;

Query Match 48.1%; Score 2621.5; DB 4; Length 507;

Best Local Similarity 90.2%; Pred. No. 9.8e-205; Matches 507; Conservative 0; Mismatches 0; Indels 55; Gaps 1;

QY 469 MSVNSEKSYDGVVDKRMQMDLRASSLVAFSTCVIKYPLRCERHGGCKTKCLAS 528
 DB 1 MSVNSEKSYDGVVDKRMQMDLRASSLVAFSTCVIKYPLRCERHGGCKTKCLAS 60
 QY 529 RPYGGMKEGACGCHLSPNSRLTFEODIERGNTDGLGDCNHSFYALNGHSSLLPSTT 588
 DB 61 RPYGGMKEGACGCHLSPNSRLTFEODIERGNTDGLGDCNHSFYALNGHSSLLPSTT 108
 QY 589 SDSTAQEGYSGRGLMDKHLDDPSDTPDLGAVSHNKKDKKGYIRRESYLKGHQLYV 648
 DB 109 -----GVIRRESYLKGHQLYV 125
 QY 649 TLTAIAVILAFVMAVFGSITVYCVCDHRKQAVVQREKELTSSRSGMSVTKLSGL 708
 DB 126 TLTAIAVILAFVMAVFGSITVYCVCDHRKQAVVQREKELTSSRSGMSVTKLSGL 185
 QY 709 PGDTOSKDPKPAIILTPLMHNGKLAIPGNTAKKLIKADQHLDLTLPTPESTPTLOQR 768
 DB 186 PGDTOSKDPKPAIILTPLMHNGKLAIPGNTAKKLIKADQHLDLTLPTPESTPTLOQR 245
 QY 769 KPSRSGREWERKQNLINACTKMPMGSEVILPTDLPLRASPSTHPIPVVVLPTTQOYGE 828
 DB 246 KPSRSGREWERKQNLINACTKMPMGSEVILPTDLPLRASPSTHPIPVVVLPTTQOYGE 305
 QY 829 YNDQPKSYAQMALQDAATLEKTIKELHLSKSNHGVNLVENLDSLPKYPQGEAL 888
 DB 306 YNDQPKSYAQMALQDAATLEKTIKELHLSKSNHGVNLVENLDSLPKYPQGEAL 365
 QY 889 GPPGASLSQGLSKRLMEHSSSYGVYKRSYPTNSLTRSHQATLLKRNNTSSNSHLS 948
 DB 366 GPPGASLSQGLSKRLMEHSSSYGVYKRSYPTNSLTRSHQATLLKRNNTSSNSHLS 425
 QY 949 RNQSPGRGNDPPAPRVDSIQVHSSQPSQAVTVRQPSLANKYNSLTSGLKRTSLKP 1008
 DB 426 RNQSPGRGNDPPAPRVDSIQVHSSQPSQAVTVRQPSLANKYNSLTSGLKRTSLKP 485
 QY 1009 DVPPKPSFAPLSTSMKPNDACT 1030
 DB 486 DVPPKPSFAPLSTSMKPNDACT 507
 RESULT 11

Q9NFY4
 ID Q9NFY4 PRELIMINARY; PRT; 1073 AA.
 AC Q9NFY4:
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Semaphorin 6D isoform 4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Qu X., Zhai Y., Wei H., Yu Y., Tang F., He F.;
 RA "Homo sapiens semaphorin 6D isoform 4 (SEMA6D.4) mRNA, complete cds";
 RT Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.
 RL EMBL, AF189429; AAM69452.1;
 DR InterPro: IPR001627; Pfam: PF01403; SMART: SM00630; SEMA: 1.
 DR SMART: SM00630; SEMA: 1.
 SQ SEQUENCE 1073 AA; 119872 MW; 7DCB4DFC5BF70F9E CRC64;

Query Match 43.4%; Score 2368; DB 4; Length 1073;

Best Local Similarity 46.0%; Pred. No. 1.5e-183; Matches 512; Conservative 159; Mismatches 308; Indels 134; Gaps 30;

QY 1 MRSEALLIYFTLL--HFAGAGFPEDSEPTISHSNTKQYVFPVGHKPGRTTQRHLD 57
 DB 1 MRSEALLIYFTLL--HFAGAGFPEDSEPTISHSNTKQYVFPVGHKPGRTTQRHLD 58
 QY 58 IOMIMNGTLYIARDDIYTVDIPTSHTEIEYCSKLTWRSRQADVDTGRMKGHKDEC 117
 DB 59 FOLMLKIDTLYIARDDIYTVDIPTSHTEIEYCSKLTWRSRQADVDTGRMKGHKDEC 118
 QY 118 HNFIVKLLKNDDALFVCGTNAFNSCNRKMDTLEPFGDEFSGMARCPYDAKANALF 177
 DB 119 HNFIVKLLKNDDALFVCGTNAFNSCNRKMDTLEPFGDEFSGMARCPYDAKANALF 178
 QY 178 ADGKLYSATVDFLAIDAVITRSLSGEPTLTIVKHSKWLKMPYVQAVDGYTFEPR 237
 DB 179 ADGKLYSATVDFLAIDAVITRSLSGEPTLTIVKHSKWLKMPYVQAVDGYTFEPR 238
 QY 238 ELAIEVNTMGKVFPRVAVQVCKNDMGSGQVLEKQMTSFLKARLNCVPGDSHFYFNILQ 297
 DB 239 ELAIEVNTMGKVFPRVAVQVCKNDMGSGQVLEKQMTSFLKARLNCVPGDSHFYFNILQ 298
 QY 298 AVTDVIRINGRDVLTATSTPYNSIPGSAVCAVMDLTASVETGRFKQKSDSTTPVP 357
 DB 299 AVTDVIRINGRDVLTATSTPYNSIPGSAVCAVMDLTASVETGRFKQKSDSTTPVP 358
 QY 358 DERVKKPPGCCAGSSSLERVATSNPPDNLNFKTHPLMDAEPVIFNRPMFLRTWVR 417
 DB 359 DERVKKPPGCCAGSSSLERVATSNPPDNLNFKTHPLMDAEPVIFNRPMFLRTWVR 418
 QY 418 YELTKIAVDTAGPYQNTTVVFLGSEKGIILKFLARIGNSGF-LNDSLFLEMSVYNSEK 476
 DB 419 YELTKIAVDTAGPYQNTTVVFLGSEKGIILKFLARIGNSGF-LNDSLFLEMSVYNSEK 476
 QY 476 CSYDGVEDKRLNGMQLDRASSSLYVAFSTCVIKYPLRCERHGGCKTKCLASDPYCGMI 536
 DB 477 CSYDGVEDKRLNGMQLDRASSSLYVAFSTCVIKYPLRCERHGGCKTKCLASDPYCGMI 536
 QY 537 XEGGACSHLSPNSRLTFEODIERGNTDGLGDCNHSFYALNGHSSLLPSTTSDTAQ 595
 DB 537 XEGGACSHLSPNSRLTFEODIERGNTDGLGDCNHSFYALNGHSSLLPSTTSDTAQ 595
 QY 596 GYSEKRG-----MLD-WKHLL-----DSP-----DSTDP 618
 DB 596 GYSEKRG-----MLD-WKHLL-----DSP-----DSTDP 618
 QY 618 YKIRGPTSMDEVSSSVTMAISPEIRPKYIDTWPRKLTSSRKFTVODDPTSDFTDP 638
 DB 618 YKIRGPTSMDEVSSSVTMAISPEIRPKYIDTWPRKLTSSRKFTVODDPTSDFTDP 638
 QY 638 LGAVSSHNDKKGYIRRESYLKGHQLYVPTTLTAIAVILAFVMAVFGSITVYCVCD-HR 677

```

Db      639  LSGT-----PKGVMEVSGESNOVMMNNVILITVFAAFVLGATLGAANYCYTDMV 691
Qy      678  RKDVAVVORKEKELTSTRSGSMSSVTKLSGLFG---DTQSKDPKPEALITPLMNGKLA 733
Db      692  RKNRKT--HKDASAGSCTDSSGSPAKLNGLPFSVPKEXQONIDSPKLYSNLTSKELP 749
Qy      734  TPENTAKMLIKADQHLDTALTPTPESTPTLQOKRPSRSGREMERNOULINACTDMP 793
Db      750  PNDITKSMVMDHRGQPEELAALEPTPESTPYLHQKTLQAMKSHSEKAGH--GASRKEHQ 807
Qy      794  MGSPLVPTDLPLASPSHPSVAVVLPITQOYQHEV-----VDPP--KMS 836
Db      808  FPPSSSEPPSPFL--SHGHLPASVLPVPAHTDYNTSFSNSNAKAEKKLQNIHPLTKSS 865
Qy      837  EVAQMALEDQAATLEYKTIKEHLSSKSPN-----HGVNLVENLDT--PPXP 882
Db      866  KRDRRSVDSRNTL--NDLLKHLNDPNSNKALMGDIQMAHQULMDPMGMEVEPPKYP 923
Qy      883  QREASLCPGASLSQGLSKRLSMHSSSY--GVYKSYPTNSLTRSHQATTLKRNNT 939
Db      924  NREASLYSPSTLPRNSPTKRVDPPTTPGVMTSLERQRYHNSQR--HSISAMPK-ML 981
Qy      940  NSNSNSHLSENQSGRD--NPPAPQVDSIQVHSSQSPGQAVTAROPSLMAYNSLT-- 996
Db      982  NSPVGVLISQPSMNRGVMPTPTGAKVDYIQ-----GTPVSYHLQPSLSQSSYTSN 1034
Qy      997  ---RSGLKRTPSLKPDVPPKPSFAPLSTMKP 1025
Db      1035  GTLPRTGLKRTPSLKPDVPPKPSFVPTQPSVRP 1067

RESULT 12
QBNFY5  QBNFY5  PRELIMINARY;  PRT;  1017 AA.
ID      QBNFY5
AC      QBNFY5;
DT      01-OCT-2002 (TtEMBLrel. 22, Created)
DT      01-OCT-2002 (TtEMBLrel. 22, Last sequence update)
DT      01-MAR-2003 (TtEMBLrel. 23, Last annotation update)
DE      Semaphorin 6D isoform 3.
GN      SEMA6D.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX      NCBI TaxID=3606;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Brain;
RA      Ou X., Wei H., Zhai Y., Yu Y., Tang F., He F.;
RT      "Homo sapiens semaphorin 6D isoform 3 (SEMA6D.3) mRNA, complete cds.";
RL      Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AF389428; AAM69451.1; -
DR      InterPro; IPR001627; Sema.
DR      Pfam; PF01403; Sema; 1.
DR      SMART; SM00630; Sema; 1.
SQ      SEQUENCE 1017 AA; 113736 MW; 4D639CEBAD9F2A0 CRC64;

Query Match 43.1%; Score 2349; DB 4; Length 1017;
Best Local Similarity 46.7%; Pred. No. 4.9e-182;
Matches 503; Conservative 155; Mismatches 302; Indels 116; Gaps 27;

```

```

Qy      1  MRSALLTYFTLL--HFAAGCFPEDESEPISSHGNTKQYVFGKPGKNTTORHLD 57
Db      1  MRVLLCAVITLLMWSQRAVSPEDDEPLNTUTVYHSRQYVPRG-RPSSNESQ-HRLD 58
Qy      58  IQMTIMNGTLYIARDHIYVDIDTSHTEIYCSKULTWKSRQADVDTCMKAKHDEC 117
Db      59  FQMLKLRDITLYINGRDQVYVNLNEMPKTEVIVNKKLTWRSRQODRENCAMKSKHDEC 118
Qy      118  HNFILVLLKKDDALFVCGTNAFNPSCNRYKADTLEPPGDFSSMACPYAKKANALF 177
Db      119  HNFILVFPVRNDEWVFCGTNAFNPMPCRYLSTLEYDGEISGLARCPDARQTNALF 178

```

```

Qy      178  ADGKLYSATVDDFLAIDAVIYRSLSGSEPLRTKVDKSKMLKPEYFVQAVDGYTYEER 237
Db      179  ADGKLYSATVADFLASDAVITRSMWDSGALKRTIKDSKWIKEPHFLHAIEGNYEYFER 238
Qy      238  ELAVENMTGKVYFPRVAOVCKNDMGSSQRYLEKQWTSFLKARLNCVPGDSHFYFNIIQ 297
Db      239  ELAVENHNLGKAVYSRVARICNDMGSSQRYLEKQWTSFLKARLNCVPGDSHFYFDVIQ 298
Qy      298  AVTDVIRINGRVVATSTPYNSTPGSAVYCYMDLIDASVFTGRFKEKSDSTMTVPV 357
Db      299  SITDIIQINGIPIVVGVTQNLNIPGSAVCAFSWDDIEKVRGKREKTDPSVTAIVP 358
Qy      358  DERVFKPRPGCCAGSSSLERVATSNPEPDNTNFKTHTPLMDAEPVSIENRPMFLATMYR 417
Db      359  EDKVRKPRPGCCAKGLAAYKTSIDFPETISFLKSHLMASVAPPIIDEPWFKTRVR 418
Qy      418  YRLTIAVDTAAGPYQNTVFLGSEKGIILFELARIGNSG-LNDSLFLEMSVYNSKR 476
Db      419  YRLTIAISVDHSAQPYQNTVIVFGSEBAGWLVKVLAK--TSPSLNDSVLLLEEIEAYNHAK 476
Qy      477  GSYDVEDEKRIKMGQDLRASSSLYVAFSTCVIKVLGRCERHKKCKTGLASRDPCGM 536
Db      477  CSANEDKRVISLQDKDHALYVAFSSCIIRIPLSRCERYGCKKSCIASRDPICGM 536
Qy      537  KEGGACSHLSPNSRLT-PEODIERGNTDGLGCHNSFVALNGHSSSLPSTTSSTAOE 595
Db      537  SQ-GSGCRVTPGMLABGYQDTEFGNTAHLGCH-----ELPSTSTPD----- 579
Qy      596  GYESRGKMLDKHLNDSPDSTPDLGAVSHNHQDKGVREBYLKHDDOLVPTLLAIIV 655
Db      580  -YKIFGG-----PTS-----GVMEVSGESNOVMMNNVILITV 612
Qy      656  ILAFTWGAVFSGITVYCVCD-HRKDVAVVORKEKELTSTRSGSMSSVTKLSGLFG--- 710
Db      613  FAFLVGAFTAGVAYCYRDMFVRKRNKT--HKDASAGSCTDSSGSPAKLNGLPFSVPK 670
Qy      711  DTQSKDPKPEALITPLMNGKLAATPGNTAKMLIKADQHLDTALTPTPESTPTLQOKRP 770
Db      671  EYQONIDSPKLYSNLTSKELPKNPDGTSVMMDHRGQPEELAALEPTPESTPYLHQKTLQ 730
Qy      771  SRGSRWERNQNLINACTCDMPMGSPVIFTLPLRASPSHPSVAVVLPITQOYQHEV- 829
Db      731  AMKSHSEKAGH--GASRKEHQPFPPSSPPSPFL--SHGHLPASVLPVPAHTDYNTSFS 786
Qy      830  -----VDPP--KMSEVQMALEDQAATLEYKTIKEHLSSKSPN----- 865
Db      787  NSNAHKAEEKLQNIHPLTKSSSKRDRRSVDSRNTL--NDLLKHLNDPNSNKALMGDI 844
Qy      866  ---HGVNLVENLDT--PPXPQREASLCPGASLSQGLSKRLSMHSSSY--GVY 916
Db      845  QMAHQULMDPMGMEVEPPKYPVNRASLYSPSTLPRNSPTKRVDPPTTPGVMTSLER 904
Qy      917  KRKYPTNSLTRSHQATTLKRNNTNSNSHLSNQSGFGD--NPPAPQVDSIQVHSSQ 975
Db      905  QRYGYNHNSQR--HSISAMPK-NLNSPVGVLISQPSMNRGVMPTPTGAKVDYIQ----- 957
Qy      976  PSCQAVTAROPSLMAYNSLT-----RSGLKRTPSLKPDVPPKPSFAPLSTMKP 1025
Db      958  ---GTPVSYHLQPSLSQSSYTSNGTLPRGLKRTPSLKPDVPPKPSFVPTQPSVRP 1011

```

```

RESULT 13
QBNFY6  QBNFY6  PRELIMINARY;  PRT;  998 AA.
ID      QBNFY6
AC      QBNFY6;
DT      01-OCT-2002 (TtEMBLrel. 22, Created)
DT      01-OCT-2002 (TtEMBLrel. 22, Last sequence update)
DT      01-MAR-2003 (TtEMBLrel. 23, Last annotation update)
DE      Semaphorin 6D isoform 2.
GN      SEMA6D.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```


Thu Oct 23 17:22:37 2003

us-09-856-681-2.rpt

Page 11

```

D6      DPN$KPAIMEDICMAHONMLDDEPG$SEVPKVPIN$EASLY$BPSTLLEN$TKRVDY  896
QY      H$SS$---GV$DYK$SYPTN$LT$R$HQA$TLKRNNTN$N$G$H$R$N$O$F$G$R$D--NEPPA  962
Db      PTTPEVEMPT$LEROG$YHKN$SSQR-H$ISAMPK-NL$N$PNCVLL$R$O$P$SM$RG$YMPPT  954
QY      PQRV$D$IOV$H$S$O$P$G$Q$AVT$R$O$P$SL$A$V$N$SLT-----R$G$K$R$T$SL$K$D$V$P$K$S$F  1016
Db      955 GAKVDYIO-----GTPV$YH$LO$P$SL$R$O$S$Y$N$G$LT$P$RT$G$K$R$T$SL$K$D$V$P$K$S$F  1007
QY      1017 APL$T$SM$K  1025
Db      1008 VPQ$T$SV$P  1016

```

Search completed: October 23, 2003, 17:12:38
Job time : 65 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 23, 2003, 17:09:07 ; Search time 17 Seconds

(without alignments)
2849.265 Million cell updates/sec

Title: US-09-856-681-2

Perfect score: 5450

Sequence: 1 MRSEALLLYTLHFAGAGF.....PPKSPFAPLSTSKPNDACT 1030

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 segs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5450	100.0	1030	1	SM6A_HUMAN
2	4346.5	79.8	888	1	SM6A_MOUSE
3	2013.5	36.9	888	1	SM6B_HUMAN
4	1997	36.6	887	1	SM6B_RAT
5	1515.5	27.8	886	1	SM6B_MOUSE
6	1515.5	27.4	930	1	SM6C_HUMAN
7	1493.5	27.4	931	1	SM6C_MOUSE
8	1491	27.4	960	1	SM6C_RAT
9	971.5	17.8	730	1	SM1A_SCHAM
10	932.5	17.1	771	1	SM1A_DROME
11	858	15.7	772	1	SM3A_RAT
12	853	15.7	772	1	SM3A_MOUSE
13	852.5	15.6	771	1	SM3A_HUMAN
14	845	15.5	712	1	SM1A_TRICF
15	845	15.5	778	1	SM1A_BRARE
16	845	15.5	860	1	SM1A_BRARE
17	839.5	15.4	772	1	SM3A_CHICK
18	830	15.2	1074	1	SM5A_HUMAN
19	826	15.2	749	1	SM5B_HUMAN
20	825.5	15.1	1077	1	SM5B_MOUSE
21	803	14.7	748	1	SM3B_MOUSE
22	803	14.7	764	1	SM3D_HUMAN
23	797	14.6	777	1	SM3D_HUMAN
24	793.5	14.6	761	1	SM3D_CHICK
25	789	14.5	712	1	SM1A_CAEEL
26	765	14.1	1093	1	SM5B_MOUSE
27	762.5	14.0	706	1	SM2A_DROME
28	762	14.0	775	1	SM3E_HUMAN
29	752	13.8	751	1	SM3C_CHICK
30	751	13.8	861	1	SM4C_MOUSE
31	750.5	13.8	697	1	SM2A_SCHGR
32	749	13.7	785	1	SM3F_HUMAN
33	744.5	13.7	785	1	SM3F_MOUSE

34	739	13.6	775	1	SM3E_MOUSE	P70275 mus musculus
35	737	13.5	751	1	SM3C_MOUSE	O62181 mus musculus
36	735.5	13.5	862	1	SM4D_HUMAN	O92854 homo sapien
37	733.5	13.5	785	1	SM3E_CHICK	O42237 gallus gall
38	733	13.4	751	1	SM3C_HUMAN	O99985 homo sapien
39	700.5	12.9	766	1	SM27_BRARE	O94151 mus musculus
40	692	12.7	834	1	SM4C_MOUSE	O64151 mus musculus
41	681	12.5	832	1	SM4B_HUMAN	O99362 homo sapien
42	671	12.3	761	1	SM4A_HUMAN	O91351 homo sapien
43	664	12.2	776	1	SM4F_RAT	O92143 ratu
44	658	12.1	777	1	SM4F_MOUSE	O92123 mus musculus
45	656	12.0	782	1	SM4B_MOUSE	O62179 mus musculus

ALIGNMENTS

RESULT 1
ID SM6A_HUMAN STANDARD; PRT; 1030 AA.
AC Q9H2E6; Q9P2H9; 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Semaphorin 6A precursor (Semaphorin VIA) (Sema VIA) (Semaphorin 6A-1) (SEMA6A-1).
DE (SEMA6A-1).
GN SEMA6A OR KIAA1368.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND INTERACTION WITH EVL.
RX MEDLINE=20564339; PubMed=10993894;
RA Klottermann A., Lutz B., Gertler F., Behl C.;
RT "The orthologous human and murine semaphorin 6A-1 proteins (SEMA6A-1/Sema6A-1) bind to the enabled/vasodilator-stimulated phosphoprotein-like protein (EVL) via a novel carboxyl-terminal tyrosine-like domain." J. Biol. Chem. 275:39647-39653(2000).
RL J. Biol. Chem. 275:39647-39653(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20181126; PubMed=10718198;
RA Nagase T., Kikuno R., Ishikawa K.-I., Hirosewa M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XVI. The complete sequences of 150 new cDNA clones from brain which code for large proteins in vitro." J. DNA Res. 7:65-73(2000).
RL DNA Res. 7:65-73(2000).
CC -!- FUNCTION: Can act as repulsive axon guidance cues. May play a role in channeling sympathetic axons into the sympathetic chains and controlling the temporal sequence of sympathetic target innervation (by similarity).
CC -!- SUBUNIT: Active as a homodimer or oligomer. Interacts with EVL.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q9H2E6-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9H2E6-2; Sequence=VSP_007113;
CC Note=No experimental confirmation available;
CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -!- SIMILARITY: Contains 1 Sema domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----

DEVELOPMENT.
 -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 -1- SIMILARITY: Contains 1 Sema domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AF030430; AAB86408.1; -
 DR MGI; MGI:1203727; Sema6a.
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR01627; Sema.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00423; PST; 1.
 DR SMART; SM00630; Sema; 1.
 KW Signal; Transmembrane; Multigene family; Neurogenesis; Glycoprotein;
 KW Developmental protein.
 FT SIGNAL 1 18
 FT CHAIN 19 888 POTENTIAL.
 FT DOMAIN 19 649 SEMAPHORIN 6A.
 FT TRANSMEM 650 670 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 671 888 POTENTIAL.
 FT DOMAIN 56 491 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 792 819 PRO-RICH.
 FT CARBOHYD 33 33 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 49 49 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 65 65 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 282 282 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 434 434 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 461 461 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 888 AA; 99075 MW; C7094FC2B345C57 CRC64;

Query Match 79.8%; Score 4346.5; DB 1; Length 888;
 Best Local Similarity 94.4%; Pred. No. 3.5e-272;
 Matches 816; Conservative 24; Mismatches 23; Indels 1; Gaps 1;

QY 1 MRSEALLVFTLTHAGAGFPEDSPISISGNTKQVFGHKGKNTQRRHLDLQ 60
 DB 1 MRPALLLCTLHCAGAGFPEDSPISISGNTKQVFGHKGKNTQRRHLDLQ 60
 QY 61 IMINAGTYIARADITVDIDTSHTEELYSKLTWMSRQADVDTCKMKGKHQECNF 120
 DB 61 IMINRRTLYVARKDHIYVIDIDTSHTEELYSKLTWMSRQADVDTCKMKGKHQECNF 120
 QY 121 IKVLLKKDDALFVCGTNAFNPSCNYYKMDTLEPGDESGMARCPYDAKANVALLPADG 180
 DB 121 IKVLLKKDDTLFVCGTNAFNPSCNYYKMDTLEPGDESGMARCPYDAKANVALLPADG 180
 QY 181 KIYKSTVDELADAVIYRSIGESPTLRTVHKDSKWLKEPYVQAVDGYIFFFREIA 240
 DB 181 KIYKSTVDELADAVIYRSIGESPTLRTVHKDSKWLKEPYVQAVDGYIFFFREIA 240
 QY 241 VEYNTMGKVPFPRVAQVCKNDMGSGQVLEKQWTSFLKARLNCVPGDSHFYFNILQAVT 300
 DB 241 VEYNTMGKVPFPRVAQVCKNDMGSGQVLEKQWTSFLKARLNCVPGDSHFYFNILQAVT 300
 QY 301 DVIRINGRDVLAFTSTPYNSIPGSVAVCYMDLADIAVTFGRFKGKSDSTWTPVDDR 360
 DB 301 DVIRINGRDVLAFTSTPYNSIPGSVAVCYMDLADIAVTFGRFKGKSDSTWTPVDDR 360
 QY 361 VKPRPGCCAGSSSERVATNSNEFPDPTLNFIKTHPLDEAVPSIENRPFRTWRYEL 420
 DB 361 VKPRPGCCAGSSSERVATNSNEFPDPTLNFIKTHPLDEAVPSIENRPFRTWRYEL 420
 QY 421 TKIAVDTAAGPYONTVFLGSEKGIILKFLARIGSGFLNDSLPLEKSVYNSEKCYSD 480
 DB 421 TKIAVDTAAGPYONTVFLGSEKGIILKFLARIGSGFLNDSLPLEKSVYNSEKCYSD 480
 QY 481 GVEDKRIKMGQDLRASSSLVAFSTCVIKVPLGRCHRGKCKTCLASNDPYCGWKIEEG 540

DB 481 GVEDKRIKMGQDLRASSSLVAFSTCVIKVPLGRCHRGKCKTCLASNDPYCGWKIEEG 540
 QY 541 ACSHLSPNSRLTPEQDIREKNTDGLDCNNSFALNGHSSLLPSTTSDSTOEYER 600
 DB 541 ACSHLSPNSRLTPEQDIREKNTDGLDCNNSFALNGHSSLLPSTTSDSTOEYER 600
 QY 601 GGWLMDKHLIDSDSDTDPGAVSSHHQDKGVIRESYLGKQDLPVTLAIAVIAFV 660
 DB 601 GGWLMDKHLIDSDSDTDPGAVSSHHQDKGVIRESYLGKQDLPVTLAIAVIAFV 660
 QY 661 MGAVFSGITVYCYCDHRKQVAVQREKELTHSRGSMSSVTKLSGLFEDTQSKDPKE 720
 DB 661 MGAVFSGITVYCYCDHRKQVAVQREKELTHSRGSMSSVTKLSGLFEDTQSKDPKE 720
 QY 721 AITPLMENGKATPGNTAKMLIKAOHLDLALPPESTPTLQOKRSPSGREMER 780
 DB 721 AITPLMENGKATPGNTAKMLIKAOHLDLALPPESTPTLQOKRSPSGREMER 780
 QY 781 ONLINCTKMPMPGSPVPTDPLRASPSPHPSVVLPTTQOGYQHEVYDQPMSE-VA 839
 DB 781 ONLINCTKMPMPGSPVPTDPLRASPSPHPSVVLPTTQOGYQHEVYDQPMSE-VA 839
 QY 840 QMALEDQATLEXYKTIKHLSSKS 863
 DB 841 QMALEDQATLEXYKTIKHLSSKS 864

RESULT 3
 SM6B_HUMAN STANDARD; PRT; 888 AA.
 ID SM6B_HUMAN
 AC Q9H3T3; Q9NRK9;
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Semaphorin 6B precursor (Semaphorin Z) (Sema Z).
 GN SEMA6B OR SEMAZ.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC Tissue=Brain;
 RA Kimura T., Ishida H.;
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=21248680; PubMed=11350127;
 RA Correa R.G., Sasahara R.M., Bengtson M.H., Katayama M.L.H.,
 RA Salim A.C.M., Brentani M.W., Sogayar M.C., de Souza S.J.,
 RA Simpson A.J.G.;
 RT "Human semaphorin 6b";
 RL Genomics 73:343-348(2001).
 CC -1- FUNCTION: MAY PLAY A ROLE IN BOTH PERIPHERAL AND CENTRAL NERVOUS
 CC SYSTEM DEVELOPMENT (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing. Named isoforms=2;
 CC Comment=Additional isoforms seem to exist;
 CC Name=1;
 CC IsoId=Q9H3T3-1; Sequence=Displayed;
 CC Name=2; Synonyms=6B.1;
 CC IsoId=Q9H3T3-2; Sequence=VSP_006044, VSP_006045;
 CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -1- SIMILARITY: Contains 1 Sema domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00423; PSI; 1.
 DR SMART; SM00630; Sema; 1.
 DR Signal; Transmembrane; Multigene family; Neurogenesis; Glycoprotein;
 KW Developmental protein.
 FT SIGNAL 1 26
 FT CHAIN 27 887
 FT DOMAIN 27 605
 FT TRANSMEM 606 626
 FT DOMAIN 627 887
 FT DOMAIN 239 549
 FT CARBOHYD 75 75
 FT CARBOHYD 156 156
 FT CARBOHYD 168 168
 FT CARBOHYD 292 292
 FT CARBOHYD 387 387
 FT CARBOHYD 442 442
 FT CARBOHYD 463 463
 FT SEQUENCE 887 AA; 95752 MW; 095433F3F202CD301 CRC64;
 Query Match 36.6%; Score 1997; DB 1; Length 887;
 Best Local Similarity 42.7%; Pred. No. 7, 1e-121;
 Matches 433; Conservative 132; Mismatches 260; Indels 190; Gaps 25;
 2 RSEALLLEFLLHFGAGFPEDESEPISSHGNYTKQYFVGHKGRNTTQ--HRLDIQ 59
 10 RPAFLFLLLLRVHGFPEDEPPPLSVAPRDYLSHYVFGSGGR.LTPAGADLNIQ 69
 60 MIMMNGLYIARDHIYTVDTIDTSHTEIYCSKULTKSROADVDTGRMKGHDECHN 119
 70 RVLAVNRTLFIDRDLNLYOVELPESTSTELRYORLTLRNSNDIDVCRMKQGECHN 129
 120 FIKYLLKXNDLAFVCGTNAFNSCRNKMOTLEPFGEFSGMACPYDAKIANVALPAD 179
 130 FVKVLLDESTLFCVGSNAFNPICANYSMDTLQLGDNISMACCPDPRIANVALPSD 189
 180 GKLVSATVDFLAIDAVIYRSLGESPTLRVGHDSKMLKEFYVQAVDYGDIYFFPRRI 239
 190 GMLFPAIVDFLAIDAVIYRSLGDRPLRTVGHDSKMEKEFYVAIVEGSHVFFPREI 249
 240 AVEYNTMGVPPRAVQVCKNMGSQRLTEKOMTSFLKARLNCVPGSHHYFILLQAV 299
 250 AMEFYLEKVVSVRAVCKNDVGSPLYLEKQWTSFLARLNCVPGSHHYFILLQAV 309
 300 TVVIRNGDVLAFTSTPYNISPGAVCAVMDLIASVFTGRFKOKSPDSTWTPVDE 359
 310 TGVVSLGRPVILAVFTSTNSIPGAVCAVPMNQVAVFEGRFRQKSPESIMTPVED 369
 360 RVPKPRPGCCAGSSSLERYATSNFPDDTLNFIKTHPLMDEAVPSIFENRPFRLTWYR 419
 370 QVPRPRPGCCAFGM--QYNASVALDEILNFVKTHPLMDEAVPSLGSHSPWIRTLIRHQ 427
 420 LTKIADVTAAGPYQNTVVFILGSEKIIILKFLAR--IGNSGLNDLFLLEMSVYNSEK 477
 428 LTRVADVAGAPWGNQITVILGSEVGTVKFLVKPNASVSGTTPSIFLEEBETYPKDDC 487
 478 ---SYGVGDKRMGMQDLBASSLVAASTCVIKYPLGRCERHGGCKTKCIASRDPYG 534
 488 GSSSSGEMGQRLSLIELDAASGLLAAPRCVVRFPVARCQVSGCMNCIGSQDPYG 547
 535 WKEGACSHLSNSRLTEQDIERNQDGLGDCHNSFYALNGHSSLLPSTTSTQ 594
 548 WARD-GSCIFLRPGTATFEQDVGASTSLGDC----- 580
 595 EGYESRGMLDMKHLDSPSTPLAVSHNQDKGVIRESYLKGHDVLEVTLLATA 654
 581 -----TGLRSLSDRAGLVSNLLVTS 604
 655 VILAFVWAGVFSGLTV--YVCCHRRKQVAVVQREKE--LTHSRGSGMSVTKL----- 705
 605 SYAAFFVAGVAVSGSVGMVFLGRERELA--RKDKVAILAHGSEAVSVSLGRGRT 662

QY 706 -SGLFQDTSKDKPKPEALITPLMHNGKLATPGNTAKMLIKADQHHDLTALPPESTPT 763
 DB 663 GTGGRGAGGPGGPPPALAPLMONG-----TKALLHGPHDLSGLPTPEQTP- 715
 QY 764 LQGRKPSR-----GSREMERONLINACTK-----DWPNGSPVITPD-- 802
 DB 716 LPQKRLFTTPHAAAGPRAWDSHALLSASASTLLLAHTRAPEQP-----VPTESG 770
 QY 803 -----LPLRASHPISVVVLPITQGGYGHVYDQPKMSEVAMALDQAAITLEYTK 856
 DB 771 PESRLCAPRSCRAHPGDPPLTP-----HASPDRRRVSAPTGPLDSSSVG-- 816
 QY 857 EHLSSKSPNNGVNLVENDLSP--PKVQREASL-----GPPGASLSQGTGSKRLEMHS 909
 DB 817 -----DLPGFWSPPATSSLRPPGHPPTALART-----HT 849
 QY 910 SSYGVYKRSYPTNSLTRSHQATTLKRNNTSSNSHLSRNQSGRQD-NPPAP 963
 DB 850 FNSG-----EAPFGHRRRRA-----PADSTHL--LPGGTGERTAPVP 887
 RESULT 5
 SM6_MOUSE
 ID SM6_MOUSE STANDARD; PRT; 886 AA.
 AC 054951;
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Semaphorin 6B precursor (Semaphorin VIB) (Sema VIB) (Semaphorin N) (Sema N).
 GN SEMA6B OR SEMAN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98027184; PubMed=9361278;
 RA Eckhardt F., Behr O., Calucci E., Yonezawa K., Nishimoto I.,
 RA Fishman M.C.;
 RT "A novel transmembrane semaphorin can bind c-src";
 RL Mol. Cell. Neurosci. 9:409-419(1997).
 CC -1- SUBUNIT: HOMODIMER. BINDS SPECIFICALLY THE SH3 DOMAIN OF THE PROTOONCOGENE C-SRC.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: DURING DEVELOPMENT IT IS EXPRESSED IN SUBREGIONS OF THE NERVOUS SYSTEM AND IS PARTICULARLY PROMINENT IN MUSCLE. IN ADULTHOOD, IT IS EXPRESSED UBQUITOUSLY.
 CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -1- SIMILARITY: Contains 1 Sema domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF036585; AAC00493.1; -
 CC MGD; MGI:102889; Sema6b.
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00423; PSI; 1.
 DR SMART; SM00630; Sema; 1.
 DR Signal; Transmembrane; Multigene family; Neurogenesis; Glycoprotein;
 KW Developmental protein.
 FT SIGNAL 1 26
 FT CHAIN 27 886
 FT DOMAIN 27 605
 FT TRANSMEM 606 626
 FT SEQUENCE 886 AA; 95752 MW; 095433F3F202CD301 CRC64;
 Query Match 36.6%; Score 1997; DB 1; Length 887;
 Best Local Similarity 42.7%; Pred. No. 7, 1e-121;
 Matches 433; Conservative 132; Mismatches 260; Indels 190; Gaps 25;
 2 RSEALLLEFLLHFGAGFPEDESEPISSHGNYTKQYFVGHKGRNTTQ--HRLDIQ 59
 10 RPAFLFLLLLRVHGFPEDEPPPLSVAPRDYLSHYVFGSGGR.LTPAGADLNIQ 69
 60 MIMMNGLYIARDHIYTVDTIDTSHTEIYCSKULTKSROADVDTGRMKGHDECHN 119
 70 RVLAVNRTLFIDRDLNLYOVELPESTSTELRYORLTLRNSNDIDVCRMKQGECHN 129
 120 FIKYLLKXNDLAFVCGTNAFNSCRNKMOTLEPFGEFSGMACPYDAKIANVALPAD 179
 130 FVKVLLDESTLFCVGSNAFNPICANYSMDTLQLGDNISMACCPDPRIANVALPSD 189
 180 GKLVSATVDFLAIDAVIYRSLGESPTLRVGHDSKMLKEFYVQAVDYGDIYFFPRRI 239
 190 GMLFPAIVDFLAIDAVIYRSLGDRPLRTVGHDSKMEKEFYVAIVEGSHVFFPREI 249
 240 AVEYNTMGVPPRAVQVCKNMGSQRLTEKOMTSFLKARLNCVPGSHHYFILLQAV 299
 250 AMEFYLEKVVSVRAVCKNDVGSPLYLEKQWTSFLARLNCVPGSHHYFILLQAV 309
 300 TVVIRNGDVLAFTSTPYNISPGAVCAVMDLIASVFTGRFKOKSPDSTWTPVDE 359
 310 TGVVSLGRPVILAVFTSTNSIPGAVCAVPMNQVAVFEGRFRQKSPESIMTPVED 369
 360 RVPKPRPGCCAGSSSLERYATSNFPDDTLNFIKTHPLMDEAVPSIFENRPFRLTWYR 419
 370 QVPRPRPGCCAFGM--QYNASVALDEILNFVKTHPLMDEAVPSLGSHSPWIRTLIRHQ 427
 420 LTKIADVTAAGPYQNTVVFILGSEKIIILKFLAR--IGNSGLNDLFLLEMSVYNSEK 477
 428 LTRVADVAGAPWGNQITVILGSEVGTVKFLVKPNASVSGTTPSIFLEEBETYPKDDC 487
 478 ---SYGVGDKRMGMQDLBASSLVAASTCVIKYPLGRCERHGGCKTKCIASRDPYG 534
 488 GSSSSGEMGQRLSLIELDAASGLLAAPRCVVRFPVARCQVSGCMNCIGSQDPYG 547
 535 WKEGACSHLSNSRLTEQDIERNQDGLGDCHNSFYALNGHSSLLPSTTSTQ 594
 548 WARD-GSCIFLRPGTATFEQDVGASTSLGDC----- 580
 595 EGYESRGMLDMKHLDSPSTPLAVSHNQDKGVIRESYLKGHDVLEVTLLATA 654
 581 -----TGLRSLSDRAGLVSNLLVTS 604
 655 VILAFVWAGVFSGLTV--YVCCHRRKQVAVVQREKE--LTHSRGSGMSVTKL----- 705
 605 SYAAFFVAGVAVSGSVGMVFLGRERELA--RKDKVAILAHGSEAVSVSLGRGRT 662

```

FT DOMAIN 627 886 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 239 549 SEMA.
FT DOMAIN 751 754 POLY-LEU.
FT CARBOHYD 75 75 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 387 387 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 886 AA; 35466 MW; E5F56D125CDA574D CRC64;

Query Match 36.6%; Score 1995; DB 1; Length 886;
Best Local Similarity 43.7%; Pred. No. 9.5e-121;
Matches 438; Conservative 132; Mismatches 257; Indels 176; Gaps 28;

QY 8 LYEFTLLHFGAG--PPDESEPTISHGNYTKQYFVGHKGRNTTOR--HRLDIOMT 62
DB 13 LSFLLLLLVGYGLFEPPEPLSVAPRYLSHYFVSSGGERLTAAGAEDLNTQRL 72
QY 63 INNGTLYIARDHTYVIDIDTSHTEIYSSKLTWKSQADVDTCMKGKHDECHNFK 122
DB 73 RVNRITLFIQDRNLQYVELEPSTSTELRYQRKLTWNSPNSDIDVCMKKGKEGECNFK 132
QY 123 VLLKKNDDALFVCGTANFNSCNYKMDLTPRGDESGMARCPYAKHVALFADGKL 182
DB 133 VLLRDESTLFVCGSNAFNPICANYSMDTLQLGDSISGMARCPYDPAKVALFEDGML 192
QY 183 YSATVDFLADAVIYRSLGESPTLRTVKHDSKMLKEPFVQADYDYIYFFEREIAYE 242
DB 193 FTLATVDFLADAVIYRSLGDRPTLRTVKHDSKMLKEPFVQADYDYIYFFEREIAYE 252
QY 243 YNTMGKVPFRAVQVCKNDMGSGQRYLEKOWTFLKARLNCSPGSHHYFNLQAVTV 302
DB 253 FNYLEKVVVSRVAVRCKNDVGSPRYLEKOWTFLKARLNCSPGSHHYFNLQAVTV 312
QY 303 IIRNGEDVIAATSTYNSIPGSVACAYMDLDAFVGTGRFKQKPSDTPWVPEBP 362
DB 313 VSIQGGPVIIAVSTBSNIPGSVACAFDMQVAAAFEGEFKQKPSDTPWVPEBP 372
QY 363 KPRPGCAGSSSLERYNTSNEFPDDTLNFKTHPLMDEAVPSIFNBPMLRTVRYRLK 422
DB 373 RPRPGCAARGM--QVNASGALDELINFKTHPLMDEAVPSIGHSPWIRLTMHQRLR 430
QY 423 IAVDTAAGPYQNTVVFVLSGEGTILKFLAR--IGNSGFINDLFLFEMSVNYSK 477
DB 431 VAVDVGAGPWGNOTIYFLGSEAGVLFVLPKPAVSAGTTPSIFLEEFETYPDRGGR 490
QY 478 SYDGEVKRIMGMQIDPASSSLVAFSTCVIKVPLGCEHKGCKTKCIASRDPYCGWK 537
DB 491 SSGGEMGQRLSLJELDASGGLLAAPRCVAVRPVAVRCQLYSGCMKICISQDPYCGMAP 550
QY 538 EGGACSHLSPNSRLTFEODIERGNTDGLGDCHNSFVALNGHSSSLPSTTSDSTAEGY 597
DB 551 D-GSCIFLRPGTSAITFEQDVSGASTSLGDC----- 580
QY 598 ESRGMLDMKHLDPSTDPPLGAVSHNHQDKGVRRESYLKCHDQVYVYTLATAVIL 657
DB 581 -----TGLRAHSLSDRAGLVSVNLVTSVA 607
QY 658 AFVWGAFFSGITV-YVCDDRKKDVAVVORKEK--LTHSRGSGSVTKLSGFGDTOS 714
DB 608 AFVWGAFFSGITV-YVCDDRKKDVAVVORKEK--LTHSRGSGSVTKLSGFGDTOS 714
QY 715 KDP-----KREAILTPIMHNGKLTATPNTAKMLIKADQHLDLTALPTPESTP 762
DB 662 TGFGGAGGAGGPGGPPEALAPIMQGM-----TKAALHGGPHLDLGLPTPTPOT 715
QY 763 TLQQRKRP-----SGSREMENQULINACTMDMPMSPVPTL-----PLRASP 811
DB 716 -LPQRKRP-----SGSREMENQULINACTMDMPMSPVPTL-----PLRASP 811
QY 812 IPSVVVLPITQGGVQHEVVDQ--KMSVAVQMALEQATLEVKTIKHEHLSKSPMHGV 868

```

```

DB 763 -PQVPAEPGP-----SRLCAPRCRASHPCDFLTPVPSDRRRVVSAPTGLDPVYG- 815
QY 869 NLVENLSDLP-PKQVQREASL-----GPPASLSQTLSKRLHSSSYGVVYKASYP 921
DB 816 -----DLPLPWPSPATSSLRRPQGPPTAAALRR-----HTFNSG-----EAP 856
QY 922 TNSLRSHQATTLKRNNTSSNSHLSNQSFGKGD-NPPAP 963
DB 857 GGRPRRPR-----PADSTHL--LPGCTERTAPFPV 886

RESULT 6
SMGC_HUMAN
ID SMGC_HUMAN STANDARD; PRT; 930 AA.
AC Q9H3T2; Q8WX19; Q8WX19; Q8WXU0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Semaphorin 6C precursor (Semaphorin Y) (Sema Y).
GN SEMA6C OR SEMAY.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Brain;
RA Kimura T.; Ishida H.;
RL Submitted (JAN1999) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
RC TISSUE=Brain;
RA Ou X.; Zhai Y.; Wei H.; Zhang C.; Xing G.; Yu Y.; Wu S.; Zhang Y.;
RA Cuiyang S.; Zhou G.; He F.;
RL Submitted (JAN-2001) to the EMBL/Genbank/DBJ databases.
CC -!- FUNCTION: MAY BE A STOP SIGNAL FOR THE DORSAL ROOT GANGLION
CC NEURONS IN THEIR TARGET AREAS, AND POSSIBLY ALSO FOR OTHER
CC NEURONS. MAY ALSO BE INVOLVED IN THE MAINTENANCE AND REMODELING OF
CC NEURONAL CONNECTIONS (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Name=1; Synonyms=Short 1;
CC Name=2; Synonyms=Short 2; Sequence=Displayed;
CC Name=3; Synonyms=Long;
CC IsoId=Q9H3T2-1; Sequence=VSP_006047;
CC IsoId=Q9H3T2-2; Sequence=VSP_006046; VSP_006047;
CC Name=3; Synonyms=Long;
CC IsoId=Q9H3T2-3; Sequence=VSP_006047;
CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -!- SIMILARITY: Contains 1 Sema domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb.slb.ch).
CC
DB EMBL; AB022434; BAB20670.1; -
DB EMBL; AF339153; AAL72098.1; -
DB EMBL; AF339153; AAL72099.1; -
DB EMBL; AF339154; AAL72100.1; -
DB Genem; HGNC:10740; SEMA6C.
DB InterPro; IPR003659; Plexin-like.
DB InterPro; IPR001627; Sema.
DB Pfam; PF01403; Sema; 1.
DB SMART; SM00423; PSI; 1.
DB SMART; SM00630; Sema; 1.
KW Signal; Transmembrane; Multigene family; Neurogenesis; Glycoprotein;
KW developmental protein; Alternative splicing.
FT SIGNAL 1 24
FT CHAIN 25 930 SEMAPHORIN 6C.

```

Query Match 27.8%; Score 1515.5; DB 1; Length 930;
 Best Local Similarity 35.8%; Pred. No. 7,7e-90;
 Matches 380; Conservative 136; Mismatches 356; Indels 189; Gaps 32;

DOMAIN 25 604 EXTRACELLULAR (POTENTIAL).
 TRANSMEM 605 625 POTENTIAL.
 DOMAIN 626 930 CYTOPLASMIC (POTENTIAL).
 DOMAIN 233 540 SEMA.
 DOMAIN 662 667 POLY-PRO.
 DOMAIN 752 755 POLY-PRO.
 CARBOHYD 70 70 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CARBOHYD 286 286 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CARBOHYD 437 437 N-LINKED (GLCNAC. . .) (POTENTIAL).
 VARSPLIC 184 223 Missing (in isoform 2).
 VARSPLIC 586 586 /FtId=VSP 006046.
 VARSPLIC 586 586 Y -> YULPFGPSPGTPSPSDMHPPOSSTLGVNTR
 (in isoform 2 and isoform 3).
 CONFLICT 125 125 /FtId=VSP 006047.
 CONFLICT 252 252 I -> V (IN REF. 2; AAL72099).
 CONFLICT 455 455 R -> K (IN REF. 1).
 SEQUENCE 930 AA; 96682 MW; 8AF8814ADB84C88 CRC64;

Db 660 -GPEPPPSKDGAVOTPOLYTFLEPPPEGVPP-----ELACLPTEBESTPELV 709
 QY 767 KRRPSRGRWERNONLINACTKMDPMWSPVPTDLELRASPSH-----IPSVVLLPIT 821
 Db 710 KHLRAAD -PWENNQNMMNA-----KEGPGSRGGHAGAPARVLVPP-- 752
 QY 822 QQGYQHEVDQPKSEVAQMALEDOATLETKIKEL---SKSPNHNVLVENLDS- 876
 Db 753 -----PPGCGGQ-----AVEVTLLEELRYLHGQPPKGAEPAPLTSR 793
 QY 877 -LPEKPVQREASIGPPASLSQTLGSKYLEHWHSSSYGVYDKRSYPTNLSLRSHQATLTK 935
 Db 794 ALPPE -PALALGGSPRRHEGSPRLDY-----PPGRCASAPPA----- 833
 QY 936 RNNNTSSNSHL-----SRNOSFGRGDNPPAPQVDSIQVHSSQSPGQAVTSRQPS-- 988
 Db 834 -RPALSAAPARLVGGGGRRLPFGSGHARAPALLTRV-----PSGGPRSYSGPGKHL 883
 QY 989 -LNAVNSLRSGIKRTPSLKPDVPPKPSF-APLSTSMKPN 1026
 Db 884 LYLGRPEGYRGRALKVDEKPOLSLKPLVGPSSRAVPPN 924

RESULT 7
 SM6C_MOUSE STANDARD; PRT; 931 AA.
 ID SM6C_MOUSE
 AC Q9WNT3;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Semaphorin 6C precursor (Semaphorin Y) (Sema Y).
 GN SEMA6C OR SEMA7.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 CX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=99160821; PubMed=10049528;
 RA Kikuchi K., Chedotal A., Hanafusa H., Ujima Y., de Castro F.,
 RA Goodman C.S., Kimura T.;
 RT "Cloning and characterization of a novel class VI semaphorin,
 semaphorin Y.";
 RL Mol. Cell. Neurosci. 13:9-23 (1999).
 CC -1- FUNCTION: MAY BE A STOP SIGNAL FOR THE DORSAL ROOT GANGLION
 CC NEURONS IN THEIR TARGET AREAS, AND POSSIBLY ALSO FOR OTHER
 CC NEURONS. MAY ALSO BE INVOLVED IN THE MAINTENANCE AND REMODELING OF
 CC NEURONAL CONNECTIONS (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -1- SIMILARITY: Contains 1 Sema domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AB013729; BAA76294.1; -;
 CC MGD; MGI:1338032; Sema6c.
 CC InterPro; IPR001627; Sema.
 CC Pfam; PF01403; Sema; 1.
 CC SMART; SM00630; Sema; 1.
 CC Signal, Transmembrane, Multigene family; Neurogenesis; Glycoprotein;
 CC developmental protein.
 CC SIGNAL 1 25
 CC CHAIN 26 931 SEMAPHORIN 6C.
 CC DOMAIN 26 905 EXTRACELLULAR (POTENTIAL).
 CC TRANSMEM 606 626 POTENTIAL.
 CC DOMAIN 627 931 CYTOPLASMIC (POTENTIAL).

```

FT  DOMAIN  234  541  SEMA.
FT  DOMAIN  663  669  POLY-PRO.
FT  DOMAIN  754  757  POLY-PRO.
FT  CARBOHYD  71  71  N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD  287  287  N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD  438  438  N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ  SEQUENCE  931 AA; 99537 MW; B0D99D54209F125 CRC64;

Query Match 27.4%; Score 1493.5; DB 1; Length 931;
Best Local Similarity 35.7%; Pred. No. 2e-88;
Matches 374; Conservative 145; Mismatches 353; Indels 177; Gaps 32;

QY  6 LLLYFTLLHAGAGGEPEDSEFISISGNYTKQYPVVGHGKGNNTQRRHLDIOMIMN 65
DB  13 LLLLSLSPQAALFPQDPTPLTSDIQAGSPSSWRGLBDDVAAL-GLDQRLTLTA 71
QY  66 GLIYIARHDIYVVDIDTSHT-EEIYCSKKLTWKSROADVDTCRMKGNHDECHNEIKYL 124
DB  72 RTLLVARDHVSPFDIQAGEEGGLVPMNKFLTRSG--DMENCAYRGLTDECYNYIRVL 129
QY  125 LKKNDDALFYCGTNAENPSCRYNTKMDLPEPGEFSGMACPYDAYAHVAFADGKLYS 184
DB  130 VPMNSQTLTACGTNSFPMCRSYGITSLOEGEELSGQACPDATQSTVAIFAEBSLYS 189
QY  185 ATVTDFLAIDAVIYRSIGSEPTLRTWKHDSKWLKEPFYQAVDYGVYFPEFEIAYEN 244
DB  190 ATADPOASDAVYRSLGPOPLRSKAYDSKWLREPHFVALHGEHVYFFREVEDA 249
QY  245 TMGVVFPVRAVYCKNDMGSGSVLEKQMTSPFKARINCSVPDSHFYENIQAVTDIR 304
DB  250 RLGRVGSRAVYCKNDMGSGSPALDHRMTSPFKRLINCSVPDSHFYFVDFVLSLTGPVN 309
QY  305 INGRDVLALFFSPYNSIPSSAICAYMDLIAVFTGRFEEKSPDSTWVPVDEVPK 364
DB  310 LHGSALFVGYFTQYNSIPSAVCAFYLDIERGFEKFEKRESLDGMATPVEDKVPSP 369
QY  365 RPGCGSSSLERATNEFPDPTLNTKCHPLMDEAVPSIFRPMFLRTMAYRFLTKIA 424
DB  370 RPSGCAVGAASFSQDLPDVLFLFKHPLIDRVPATQIP-LITLSALLTQVA 428
QY  425 VDTAAGPYQNTVTVFLGSEKGIILKFLARIGNSGFLNLSLEMSVYNSKCSYDQVED 484
DB  429 VDGAGPHRNTTVFLGSDNGTVLKVLP-CCGSLGSEPIVLEHIDAYSHARCS--GKRS 485
QY  485 ----KRMGQDLBASSLYVAFTCYIKVPLGRCEHNGCKCTCLASRPYCGWIKEG 540
DB  486 PRAARITIGELDEGRHLEFVAPGCGCYVLSLRCARHGAQCRCLASLDPYCGWHSRG 545
QY  541 ACSHLSPNSRLTEFQDIERGNTDGLRCHNSFVALNGHSSSLPSTTSDSTAQGYE-S 599
DB  546 CMSIRGP-----GGTD-----VDLTGNS-----TEHGDDQDGAITS 578
QY  600 RGMGLDMKHLIDSPDSTDLGAVASHNHQDKGVIRSYLKGHDQVPTLLATAVILAF 659
DB  579 QSGGDSAY-----GVRDLSPASASRSIPILLACVAAAF 615
QY  660 VMGAVSEGIYVYCVCDHRKDVAVVQKKEKELTHSRGNSSVYKLSGLGDPQSKPKP 719
DB  616 ALGASVGLIVSCAC--RRAN-----RRRSKDIETPGJPRPLSLSLARLGGGPEPPPP 669
QY  720 ---EAILTPLMNGKLTATPGNTAKMLKADQHNLDLALPPESTPTLQGRKPSRSGRE 776
DB  670 KDGDAQTPQLYTFLPLPPDGSGSP-----ELACLPPETTPPLPYKHLRASGG-P 719
QY  777 WERNQNLINACT-KDMPMGSPVPIPTDLF--LRASPHLPSVVVLPTQO--GYQH-- 827
DB  720 WEMNGNNGNASEGGRPRGCGSAGAPRVIVAPPBGCGAVETTLLELLRYLHGP 779
QY  828 -----EYVQPKMSEVAQMALEQOATLEKTKIKELLSKSPHGNVLENTL 874
DB  780 QPFRKSGEPLASAPFTSRPPASEGASLFDV-----SSMPMDGVPPL-RL 824
QY  875 DSLPPK-----VPOREASIGPPGALISQGLSKRLEMHSSSYGVYDKRSYPTNSILTRSHQ 930

```

```

DB  825 D-VPEGRKAAPSGRPLASAPRLG-VGGSRL-----PFPT-----HR 862
QY  931 ATTLKRNNTNSNSHLSRNSFGGNDPPAPQVDS--IQVHSSQEGCAVTSRQS 988
DB  863 A-----PFGILTRVPSGPARVSGGPRHLLYIGR-PE 894
QY  989 LNANSLTRSGUKRTPTSLKPDV--PPKPS 1015
DB  895 GHRGSLKRVKDVKSPSPKPLASPPQPA 923

RESULT 8
SMC_RAT ID SMC_RAT STANDARD; PRT; 960 AA.
AC Q9WTU3; Q9WTU6;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Semaphorin 6C precursor (Semaphorin Y) (Sema Y).
GN SEMA6C OR SEMAY.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_Taxid=10116;
RN [1]
RP SOURCE FROM N.A. (ISOPFORMS SEMA Y-L AND SEMA Y-S).
RC STRAIN=Sprague-Dawley; TISSUE=Muscle;
RA MEDLINE=99160821; PubMed=10049528;
RA Kikuchi K., Chedotal A., Hanafusa H., Ujimasa Y., de Castro F.,
RA Goodman C.S., Kimura T.;
RT "Cloning and characterization of a novel class VI semaphorin,
RT semaphorin Y."
RL Mol Cell. Neurosci. 13:9-23 (1999).
CC - FUNCTION: SHOWS GROWTH CONE COLLAPSING ACTIVITY ON DORSAL ROOT
CC GANGLION (DRG) NEURONS IN VITRO. MAY BE A STOP SIGNAL FOR THE DRG
CC NEURONS IN THEIR TARGET AREAS, AND POSSIBLY ALSO FOR OTHER
CC NEURONS. MAY ALSO BE INVOLVED IN THE MAINTENANCE AND REMODELING OF
CC NEURONAL CONNECTIONS.
CC - SUBCELLULAR LOCATION: Type I membrane protein.
CC - ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Sema Y-L;
CC IsoId=Q9WTL3-1; Sequence=displayed;
CC Name=Sema Y-S;
CC IsoId=Q9WTL3-2; Sequence=VSP_006048;
CC - TISSUE SPECIFICITY: EXPRESSED IN MANY REGIONS OF THE DEVELOPING
CC NERVOUS SYSTEM, PROBABLY IN NEURONS AND THEIR PRECURSORS, BUT ALSO
CC IN NONNEURAL TISSUE SUCH AS IMMATURE MUSCLE AND DERMIS. IN ADULT,
CC STRONG EXPRESSION IN THE SKELETAL MUSCLE AND MODERATE EXPRESSION
CC IN THE BRAIN, WHERE CEREBELLUM SHOWS THE HIGHEST EXPRESSION. ALSO
CC EXPRESSED IN ALMOST ALL AREAS OF THE CNS.
CC - DEVELOPMENTAL STAGE: DETECTED AT E12 AND FOUND AT MARKEDLY
CC INCREASED LEVELS AT E15 AND E18 IN BOTH THE HEAD AND THE BODY. AT
CC BIRTH THE LEVEL DECREASES SIGNIFICANTLY.
CC - SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC - SIMILARITY: Contains 1 Sema domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.sib.ch).
CC
CC EMBL; AB000817; BAA76293.2; -
CC EMBL; AB014074; BAA76295.1; -
CC InterPro; IPR001627; Sema.
CC Pfam; PF01403; Sema; 1.
CC SMART; SM00630; Sema; 1.
CC Signal; Transmembrane; Multigene family; Neurogenesis; Glycoprotein;
CC Developmental protein; Alternative splicing.
KW

```

FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 960 SEMAPHORIN 6C.
 FT DOMAIN 24 635 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 636 656 POTENTIAL.
 FT DOMAIN 657 960 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 234 541 SEMA.
 FT DOMAIN 693 699 POLY-PRO.
 FT DOMAIN 783 786 POLY-PRO.
 FT CARBOHYD 69 69 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 285 285 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 436 436 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 586 617 Missing (in isoform Sema Y-S).
 FT /FtId=VSP 006048.
 SQ SEQUENCE 960 AA; 102610 MW; C882933C607E6086 CRC64;

Query Match 27.4%; Score 1491; DB 1; Length 960;
 Best Local Similarity 35.1%; Pred. No. 3, 1e-88;
 Matches 374; Conservative 149; Mismatches 363; Indels 180; Gaps 32.

QY 6 LLLVYTLHPAGAPDESEPISSHGNYTKQYVYVGHKPGKRNTPQHRRLDIQIMIMIN 65
 DB LLLLLLSPQATAPFPPDPIPLITSDLOGTSPSSMFRLEDDAVAAEL-GIDPQRFLLIN 69
 QY 66 GTLYIARDHIYVDIDTSHT-EEIYCSKSLTWKSRQADVDTGRMKGKHDECHNFIKYL 124
 DB 70 RTLLVARDHVSFDLQAESEGEGLVPRKFLTWRSQ--DMENCAVRGKLTDECYVIRVL 127
 QY 125 LKANDDALVCGGTNAFNSCRNYKNDLEPPGDESSGMAKPCYDAKANVALFADGKXYS 184
 DB 128 VFWDSQTLACGTNFSFVCRSYGITSLOEGEELSGQRCFDTQSTVAISAGSLYS 187
 QY 185 ATVTPEFLDAIVYISLGEFPLRTYKHSKWLKEPYEQAQADYDYIYFFREIAYEYN 244
 DB 188 ATADPQSDAVAVYSLGQPLRSKAKYDSKWLREPHFYALHEHGDHYIFFREYVSDA 247
 QY 245 TWGKVFPRVAVOYCKNDMGSGRVLEKOWTSFLKARLNCVSDSHFFENIIQAVTVIR 304
 DB 248 RIGRQYFQSVARVACRWDGSGRALDRHMTSFLKLKNCVSDSTFYVDVLOSLTPVN 307
 QY 305 INGRDVIALESTFPIINSIGSAVCAVMDLIDSVTGRKREKSDSTWTPVDRPVRPK 364
 DB 308 LHGRSALFQVFTTQNSIPGSAVCAFYLDIERGFEKREKRSIDGATVPVSEDKVBP 367
 QY 365 RRGCGAGSSLERVATNSNEFPDPTNFIKTHPLMEAVPSINRPFELTMTWRYLTKIA 424
 DB 368 RRGSCGAGAAALFSSQDLPDDVLLFIQHPILDAVPAHPHQP-LLTLSRALLTYVA 426
 QY 425 VDTAAGPYQNTVFLGSEKGIILKELARIGNSGLNDSLPLEMSVYNSEKCSYDGVED 484
 DB 427 VDGMPHNTVYLFSGNDGTLYKLP-GGQSLGPEPILLEIDAYSHARCS--GKRS 483
 QY 485 ---KRMQDIDRASSLYVAFSTVIAKPLGRCHRGKCKKTCLASDDPYCGWKKEG 540
 DB 484 PRAARIIGLEHDTBEGHRLFAVAPPGCIYVLSRCRHAQCCQCLASIDPYCGWHFRFG 543
 QY 541 ACHSLSPNRILTEODIE-RGNTDG--GDCHNSFVALN---GHSSSLP-----PSITT 588
 DB 544 CVNIRPBG---TVVDLTGNOESMEHGDQDGAIGSGSGPDSAYVLLGPSPETPS 598
 QY 589 SDSTAOEYVESRGMLDMKHLIDSPDTPLAGVSSHNQDKKGVITRESYLKGDQLVPY 648
 DB 599 SPEDAIRPGQS-----STLGA-----HTGGVARDLSPASRSRIP 634
 QY 649 TLLIAVILLAFVGVANFSGITVYCVCDHRKDVAVVQRKKEKLTHERRSGMSVTLSGI 708
 DB 635 PLLIACVAAPFALGASVGLVSCAC--RRAN---RRSKDIEPGLPPPLRLSLRL 688
 QY 709 FGPTGSDPKP---EALITPLMNGKLTAGNTAKMLIKADQHLLDALTALPTESTPTIQ 765
 DB 689 HGGGPEPPPPKGGDAQTQLYTTPRPPGSGSP-----ELACLPTEETTPELP 739
 QY 766 QKRPSPGSRERNQTLINACTKMPGSPVIFPTDLP--LRASDHPISVVVLPITO 822

DB 740 VKHLRASGG-PWMNQNNNASSBPGPRRCGSAAGPAPRVIVRPPPGCGQGEVEYTTL 798
 QY 823 Q---GYQH-----EYVDQPMSEVAQALDQAATLEYKTIKEHLSSKSP 864
 DB 799 EELLRYLHGPQPPRKSGPEPLASAPFTSRPPASEGALFYD-----SSPMP 844
 QY 865 NHGVNLVENIDSLP-----KVPRBASLPGFASLSOTGLSKLEMHSSSYG 913
 DB 845 R-----DCVPEFLRDVPPDGKRAAPSGRPLSAPAPRLGVSG--SRRL----- 885
 QY 914 VDYKRSYPTNSILRSHQATTLKNNNTSSNSHLSRQSGFGDNPAPAPQVDS--IQV 971
 DB 886 ---PPT-----PRA-----PFGILTIVPSGGGRSR 907

RESULT 9

ID SMA_SCHAM STANDARD; PRT; 730 AA.

AC Q26473;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Semaphorin 1A precursor (Semaphorin-I) (Sema I) (Fasciclin IV).
 NM SEMA-1A OR FAS4.
 OS Schistocerca americana (American grasshopper).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Preyngota;
 OC Neoptera; Orthopteroidea; Orthoptera; Cellifera; Acridomorpha;
 OC Acridoidea; Acrididae; Cyrtacanthacridinae; Schistocerca.
 OX NCBI_Taxid=7009;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=93040225; PubMed=1418998;
 RA Kolodkin A.L., Matthes D.J., O'Connor T.P., Patel N.H., Admon A.,
 RA Bentley D., Goodman C.S.;
 RT "Fasciclin IV: sequence, expression, and function during growth cone
 RT guidance in the grasshopper embryo";
 RT Neuron 9:831-845(1992).
 CC - FUNCTION: PLAY A ROLE IN GROWTH CONES GUIDANCE.
 CC - SUBCELLULAR LOCATION: Type I membrane protein.
 CC - TISSUE SPECIFICITY: DYNAMICALLY EXPRESSED ON A SUBSET OF AXON
 CC PATHWAYS IN THE DEVELOPING CNS AND ON CIRCOMFERENTIAL BANDS OF
 CC EPITHELIAL CELLS IN DEVELOPING LIMB BUDS.
 CC - SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC - SIMILARITY: Contains 1 Sema domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL; L00709; AAA29808.1; -
 CC PIR; JH0798; JH0798.
 CC InterPro; IPR003659; Plexin-like.
 CC InterPro; IPR002165; Plexin_repeat.
 CC InterPro; IPR001627; Sema.
 CC Pfam; PF01437; PSI; 1.
 CC DR Pfam; PF01403; Sema; 1.
 CC SMART; SM00423; PSI; 1.
 CC SMART; SM00630; Sema; 1.
 CC Signal; Developmental protein; Transmembrane; Glycoprotein;
 CC Neurogenesis.
 KW Neurogenesis.
 FT SIGNAL 1 20
 FT CHAIN 21 730 SEMAPHORIN 1A.
 FT DOMAIN 21 630 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 631 651 POTENTIAL.
 FT DOMAIN 652 730 CYTOPLASMIC (POTENTIAL).

```

FT DOMAIN 215 515 SEMA.
FT DOMAIN 217 222 POLY-RHE.
FT CARBOHYD 44 44 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 71 71 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 163 163 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 267 267 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 360 360 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 539 539 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 730 AA: 81214 MW: 6065594678537F CRC64;

Query Match 17.8%; Score 971.5; DB 1; Length 730;
Best Local Similarity 34.4%; Pred. No. 5,9e-55;
Matches 244; Conservative 124; Mismatches 224; Indels 117; Gaps 29;

8 LFTLLHFAAGFPDESEPISSHGVYTCQPYFVGHKGRNTQK-----HRLDIOMI 61
11 LFWVAH--AAAVNDVSP-----KMYVOF-----GEERVQFLNGSHKHDFKLL 54
62 MINNGTLYIAARDHIYTVDD--TSHTTEIYCSKLTWRSQADVTCRMKGKDECHN 119
55 EKDNHSLVAGARIYVNTSLRDLTFTEQ-----RIEMHSSGAHRELCTLKGSDEDCN 109
120 FIKVLLKNDKDALFVCGTNAFNSCKNYKMDLEPGD-----EFGMARCPIDAHNAV 174
110 YIRVLAKIDDDRVLLICGTNAKPLCRHYALKD-----GDYVVEKEVGRGLCPDPDHNSI 165
175 ALFADGKLSATVTDPLAIDAVYRSLGESPTLRVTKHDSKMLKEPYQAVDYGDIYF 234
166 AITSEGLISATVADSSGNDPLIYRG-----PLRTERSDUKQNAINFVNTMYNDFIIF 220
235 FFERIAVEYNTWKRVVPRVAQYCKNDMGSGQVLEKQWTSFLKARLNSVPGDSHFYFN 294
221 FFERIAVEYNTWKRVVPRVAQYCKNDMGSGQVLEKQWTSFLKARLNSVPGDSHFYFN 279
295 ILQAVDVIRIN-GRDV---VLATFSPYNSIGSVAACADMIDIASVFTGRKEDKSPD 350
280 EIGSTSDIIEGNYGGQVEKLIYGVFTTPVNSIGSVAACAFSMKSIIESFPGPEKQETNM 339
351 STWTPVDEDEVKPRPGCCAGSSSLERVATSNFEPDITLFIKTHPLMDAAPSIFNRPW 410
340 SNMLAVSLKVPERRPGQCND-----SRLPLDVSVNFVSHLMDAFAVAFTRPI 391
411 FLRTMVEYRLTKIADV---TAAGPYQNTVFLGSEKGIILFLARIGNSGFLN----- 461
392 LIRSLQYRFTKIAVDQVWRTPDG--KAYDVLFIQTDDGKVIKAL-----NSASFDSDDY 445
462 DSLFLEMYSYNSKCYGVDEKRIKMGQIDASSLLVAFSTCVIKVYLGRC--ERRG 519
446 DSVVIEELQVLP-----PGVPYKNIYVVRMDGDSKLVVVSODEITAIKLRGCSDKIT 499
520 KCKKTCTASRPYCGMIKEGAGCSHL-SPN---SRLTFEODIERGNTDGLGDCNHFVAL 575
500 NCRB-CVSLDDPYCAMNVELKCTAVSPMSACKRFIONISLGEHKACGGRPQETIV- 557
576 NGRSSSLLPSTTSSTDA-----QEGYSSRGGMIDWKLLDSDPSTDLPLAVASHNQ 628
558 ----ASVPYTOPTKSSGDPVSHIQAFEPF---IDNEIIVIGVDDSNVPIPTLAEINHA 610
629 DKGVIVESTLKGHDQVLPV---TL-LAI-----AVLIAFWGAVES 666
611 GSK-----LPSSQEKPLPYTAFTLTALVTSCLAGLVGFIISGFLS 652

```

```

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=94094332; PubMed=8269517;
RT Kolodkin A.L., Matthews D.J., Goodman C.S.;
RT "The semaphorin genes encode a family of transmembrane and secreted
RT growth cone guidance molecules."
RT Cell 75:1389-1399(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amaratilake P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA Geotze R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blake J.R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktiroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher C., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Gerhart W.M., Glasser K.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Harris M.,
RA Goddek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.D., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lascko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Mostrel A.,
RA Mount S.M., Moy M., Murphy L., Murthy D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson K., Skupski M.P., Smith T.,
RA Spier E.C., Spradling A.C., Stapleton W., Strong R., Sun E.,
RA Sytkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Maesman D.A., Weinstein G.M., Weisenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye U., Yeh R.-F., Zaveri J.S., Zhan W., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RT Science 287:2185-2195(2000).
CC -1- FUNCTION: PLAY A ROLE IN GROWTH CONES GUIDANCE.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: EXPRESSED BY SUBSETS OF NEURONS AND MUSCLES.
CC -1- DEVELOPMENTAL STAGE: EXPRESSION BEGINS AROUND STAGE 10, PRIMARILY
CC IN THE DEVELOPING CNS. IN STAGE 16 EMBRYOS, IT IS EXPRESSED AT
CC HIGHEST LEVELS THROUGHOUT THE CNS, AND WEAK EXPRESSION IS SEEN IN
CC PORTIONS OF THE PERIPHERAL NERVOUS SYSTEM, MOST CLEARLY IN THE
CC LATERAL SENSORY CLUSTERS.
CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -1- SIMILARITY: Contains 1 Sema domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L26082; AAA68789.1; -

```


FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 591 591 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 772 AA; 88808 MW; 240907812FE9F2D2 CRC64;

Query Match 15.7%; Score 858; DB 1; Length 772;

Best Local Similarity 32.3%; Pred. No. 1,3e-47;
 Matches 209; Conservative 104; Mismatches 251; Indels 84; Gaps 20;

QY 47 GRNTTQHRIDIMIMINGT-----LYIARDHIYVVDIDT 83
 Db 25 GKXNVPLKLSYKEMLESNVITFENGLANSSVHTFLDERGRSLVYGAADHIFSNL-- 82
 QY 84 SHRELYCSKLTWKSKQADVDTCRMKGRK-KDECHNFTVLLKDDALFVCGTAPNP 142
 Db 83 ---VNKDFGKIYWPVSYTRRDECKWAGKCIKRECNFIVLKYAQTHLYAGGTAFHP 139
 QY 143 SCR-----NYKMDTLEPFGEF--SGMRCFYDAKHANVALFADGKLYSATVDTFALDA 195
 Db 140 ICTYIEVGHHPEDNIFKLDQSHENGKSPYDPKLLTSLIDGELYSGTADPMGRDP 199
 QY 196 VIYRSGEPTLTATVKDSKMLKEPVFOA-----VDYGYIYFPRRIAYEVNTMGK 248
 Db 200 AIFRTIGHHPIRTEQDSRWLNDPPIAHLLPESDNPEDDKVYFEFFRENALDGHSGK 259
 QY 249 VVEPRVAQVCKNDMGSGQVLEKQWTSFLKARLNCVPG---DSHFYFIILQAVTDVIR 304
 Db 260 ATHARIGQICKNPFPG-HRSLVKNWTFELFARICSPGNGSIDTHP---DELQVFL 313
 QY 305 INGRD-----VLATFTPYRSIPGSAVCAYMDLISVFTGRKEOKSPDSTWTPVDER 360
 Db 314 MNSKDPKNPIYVGFSTSSNIFGSAVCYMSDVARVFLGPAHRGRGPNYQWVPY-QGR 372
 QY 361 VVPRPRCCGSSSLERVATNSPEPDTLNFITHPMDAVSIFPRPELFTMYRKL 420
 Db 373 VVPRPRCTCP-SKTFGGFSDTKDLPDVITFASHHPMYPVPINRPMITDYNQF 431
 QY 421 TKIADVDAAPYQNHVTFVFGSEKGLTKFLARLNGSGFNLDSLFLEMSVNSSEKSYD 480
 Db 432 TQIVDVRVADBDQDYMTIGTGVYLVKAVSPKEMHLEVLLEMYVFR----- 484
 QY 481 GVEDKRIWQMDLRASSSLVYAFSTCVIKVPLGCEHSGKCKTCLASRDPYCGWKEGG 540
 Db 485 --PPTISAMELSTKQOOLYIGSTAGVAQLPLHRCDIYGACAECCCLARDPYCAM--DGS 540
 QY 541 ACSHLSNS-RLTFEQIERGNDGLGDC-----HNSFVALNHS--SSLPEPTTSDST 592
 Db 541 SCSRYFTFARRRTRRDIRNG--DPLTHCSDLOHHDNH--HGSLEERLIYGVENSSTP 595
 QY 593 AOEYEGSGMDLWKHLSDPSTDPGLAVSSHNDKQKVIRESYLK 640
 Db 596 LECSPKQKRALVWQFORNEDRKEEL-RVGDHILIRTEGILLRSLOK 642

RESULT 12
 SM3A MOUSE STANDARD; PRT; 772 AA.
 ID O086E5; 062180; 062215;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Semaphorin 3A precursor (Semaphorin III) (Sema III) (Semaphorin D) (Sema D).
 GN SEMA3A OR SEMAD OR SEMD.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NMRI; TISSUE=Embryo;
 RX MEDLINE=95267431; PubMed=7748561;
 RA Pleschel A.W., Adams R.H., Betz H.;

RT "Murtine semaphorin D/collapsin is a member of a diverse gene family
 RT and creates domains inhibitory for axonal extension.";
 RT Neuron 14:941-948(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97470885; PubMed=9311345;
 RA Taniguchi M., Yuasa S., Fujisawa H., Naruse I., Saga S., Mishina M.,
 RA Yagi T.;
 RT "Disruption of semaphorin III/D gene causes severe abnormality in
 RT peripheral nerve projection.";
 RT Neuron 19:519-530(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Kimura T., Fishman M.C.;
 RT "cDNA sequence of mouse collapsin/semaphorin III.";
 RT Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 107-772 FROM N.A.
 RC TISSUE=Fetal brain;
 RX MEDLINE=95267432; PubMed=7748562;
 RA Messersmith E.K., Leonardo E.D., Shatz C.J., Tessier-Lavigne M.,
 RA Goodman C.S., Kolodkin A.L.;
 RT "Semaphorin III can function as a selective chemorepellent to pattern
 RT sensory projections in the spinal cord.";
 RT Neuron 14:949-959(1995).
 CC -1- FUNCTION: PLAY A ROLE IN GROWTH CONES GUIDANCE. MAY FUNCTION TO
 CC PATTERN SENSORY PROJECTIONS BY SELECTIVELY REPELLING AXONS THAT
 CC NORMALLY TERMINATE DORSALLY.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED EARLY IN EMBRYONIC DEVELOPMENT
 CC (E11) IN DISTINCT REGIONS OF THE NEUROECTODERM AND MESODERM.
 CC EXPRESSION BECAME MORE EXTENSIVE AT LATER STAGES.
 CC -1- DOMAIN: STRONG BINDING TO NEUROFILIN IS MEDIATED BY THE CARBOXY
 CC THIRD OF THE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X85983; CAA59985.1; -
 CC EMBL: D85028; BAA19773.1; -
 CC EMBL: LA1541; AAL77611.1; -
 CC EMBL: LA0484; AAA73934.1; -
 CC PIR: I48747; I48747.
 CC PIR: I58169; I58169.
 CC MGI: MGI:107558; Sema3a.
 DR InterPro: IPR007110; Ig-like.
 DR InterPro: IPR003599; Ig.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003659; Plexin-like.
 DR InterPro: IPR001627; Sema.
 DR Pfam: PF00047; Ig_1.
 DR Pfam: PF01403; Sema; 1.
 DR SMART: SM00409; Ig_1.
 DR SMART: SM00423; PSI; 1.
 DR SMART: SM00630; Sema; 1.
 DR PROSITE: PS50835; Ig_Like; 1.
 DR Signal: Immunoglobulin domain; Multigene family; Neurogenesis;
 KM Developmental protein; Glycoprotein.
 FT SIGNAL 1 20
 FT CHAIN 21 772
 FT DOMAIN 240 538
 FT DOMAIN 579 665
 FT DOMAIN 728 770
 FT DISUFID 650 723
 FT CARBOHYD 53 53
 N-LINKED (GLCNAC. . .) (POTENTIAL).

```

FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 591 591 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 193 193 D -> N (IN REF. 4).
FT CONFLICT 207 207 H -> D (IN REF. 1).
FT CONFLICT 253 253 D -> G (IN REF. 1).
FT CONFLICT 352 352 F -> L (IN REF. 4).
FT CONFLICT 403 403 A -> G (IN REF. 1).
FT CONFLICT 571 571 OH -> ED (IN REF. 1).
FT CONFLICT 616 620 EDKKE -> RRSKR (IN REF. 1).
FT CONFLICT 623 623 R -> K (IN REF. 4).
SQ SEQUENCE 772 AA, 88799 MW, E89A08528B10AEC3 CRC64;

Query Match 15.7%; Score 853; DB 1; Length 772;
Best Local Similarity 32.2%; Pred. No. 2.7e-47;
Matches 208; Conservative 102; Mismatches 257; Indels 78; Gaps 18;

QY 47 GRNTQRHRLDQIMINMG-----LYIARDHIYVDIT 63
DB 25 GKNVPRKLKSYKEMLESNNVITFENGLANSSYHTEFLDEERSRLYVGAKHIFSENL-- 82
QY 84 SHTEIYCSKUTMKSRQADVTCRMKGK-KDECHNFIKVLKKNDALFVCGTNAFNP 142
DB 83 ---VNIKFOKIVWVSTRBECKMAGKDIKECANFKVLEAVNQHLYACGIGAHFP 139
QY 143 SCR-----NYKMDLLEPFGEF--SGMACPYDAKHANVALFADGKYSATVDFLAIDA 195
DB 140 ICTYIEVGHHPEDNI.FKLQDSHFENGGRKSPYDPKLTLSLIDDELXSGTADPMGDF 199
QY 196 VYRSLGSEPTLRKYHDSKMLKEPYVOA-----VDYGYITFFREIAYENVMKG 248
DB 200 ALFRLGHHPIRTQHDGRMLNDPRFISAHILPESDNEDEDKVFFRENIDGESHGK 259
QY 249 VVEPVAQYCKNDMGSCVLEKQWTSFLKARLNSVPG---DSHFYFNILQAVTVYR 304
DB 260 ATHAIGQICRNDPFG-HRSLVNMKTTPLKARLISGVEPNIDHIF-----DELQDVF 313
QY 305 INGRD---VVIATSTPYNSIPSGAVCAVMDLIASVTFGEKQKSPDSTWTFVDPDR 360
DB 314 NMSKDPKPIYGVFTTSSNIFKGSAGVMSVDRVPLGYPARHDGPNYQWVY-QGR 372
QY 361 VKPRPGCCAGSSSLERYATSNRPDDTLNFKTHPLNDEAPSP.FNRPWFKTRVYRL 420
DB 373 VYPRPRGTCP-SKTGFGFSTDLDPDVTTFARSHAMNPVFPNNRIMKTIVNQF 431
QY 421 TKIADVTAGPYQNTVPLFGEKGIILFKLARGSGFNLSTLEKSYVNSSEKSYD 480
DB 432 TQIVVDVDAEDGQYDWFIQTVGLKVSVPRKETHWDLDEVLLEKTVR----- 484
QY 481 GVEDKRIKMGQIDRASSSLVYAFSTCVIKVPLGRCEHSGKCKKTCIASRDPYCGMKSG 540
DB 485 --EPTTISAMELSTKQOQLYISTAGVACLPLHRCDIYKACACCCCLARDPYCAW--DGS 540
QY 541 ACSHLSFNS-RLTFEDDIERGNTDGLGDCHNSFVALNGSSSL---LPSTTSSTAOE 595
DB 541 SCGRYFPFAKRRTRKODIRNG--DPLTHOSDIQHDHNGHGFLEERITIVGENSSTFLEC 598
QY 596 GYESRGMDMGLDSDPDTPLGAVSSHNQDKGVIRRESYLK 640
DB 599 SPKSGRALVYVQFORNEDRKEEI-RMGDHIIITREGILLRLSLQK 642

RESULT 13
SMA3A HUMAN STANDARD; PRT; 771 AA.
ID 014563;
AC 30-MAY-2000 (rel. 39, Created)
DT 30-MAY-2000 (rel. 39, Last sequence update)
DT 15-SEP-2003 (rel. 42, Last annotation update)
DE Semaphorin 3A precursor (Semaphorin III) (Sema III).
GN SEMA3A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

```

```

OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RC TISSUE=fetal brain;
RX MEDLINE=94094332; PubMed=8269517;
RA Kolodkin A.L., Matthes D.J., Goodman C.S.;
RT "The semaphorin genes encode a family of transmembrane and secreted
RL growth cone guidance molecules."
RN Cell 75:1389-1399 (1993).
RN [2]
RP SEQUENCE OF 1-37 FROM N.A.
RA Moesner J., Mink P., Hinds K., Strommatt C.;
RL Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE OF 39-182 FROM N.A.
RA Rolfing T., Tin-Wollam A.M., Duckels G.;
RL Submitting (JUN-1998) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: INDUCES THE COLLAPSE AND PARALYSIS OF NEURONAL GROWTH
CC CONES. COULD SERVE AS A LIGAND THAT GUIDES SPECIFIC GROWTH CONES
CC BY A MOTILITY-INHIBITING MECHANISM. BINDS TO THE COMPLEX
CC NEUROFILIN-1/PLEXIN-1 (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: secreted (By similarity).
CC -1- DOMAIN: STRONG BINDING TO NEUROPILIN IS MEDIATED BY THE CARBOXY
CC THIRD OF THE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, L26081, AAA65938.1, -.
DR EMBL, AC004451, -. NOT ANNOTATED. CDS.
DR EMBL, AC004848, AAC78622.1, -.
DR PIR, D49423, D49423.
DR Genew: HGNC:10723; SEMA3A.
DR MIM, 603961;
DR GO, GO:0005576; C:extracellular; TAS.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003066; Ig_MHC.
DR InterPro: IPR003659; plexin-like.
DR InterPro: IPR001627; Sema.
DR Pfam, PF00047; Ig_1.
DR Pfam, PF01403; Sema; 1.
DR SMART, SM00409; Ig; 1.
DR SMART, SM00423; PSI; 1.
DR SMART, SM00630; Sema; 1.
DR PROSITE, PS50835; IG-LIKE; 1.
KW Signal; immunoglobulin domain; Multigene family; Neurogenesis;
KW Developmental protein; Glycoprotein.
FT SIGNAL 1 20
FT CHAIN 21 771 SEMAPHORIN 3A.
FT DOMAIN 580 638 SEMA.
FT DOMAIN 727 769 ARG/IXS-RICH (BASIC).
FT DISULFID 649 722
FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 590 590 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 771 AA, 88889 MW, 9985F8D3E8ED8456 CRC64;

Query Match 15.6%; Score 852.5; DB 1; Length 771;
Best Local Similarity 32.1%; Pred. No. 2.9e-47;
Matches 208; Conservative 106; Mismatches 250; Indels 83; Gaps 20;

QY 44 HKPGNTQRHRLDQIMINMG-----LYIARDHIYVD 80

```

```

Db 22 YONGKUNVPRKLKSYKEMLEBNNVTFPENGJLANSSSSYHTFLDEBRSRLVYGAKHIFSPD 81
Qy 81 IDSHTEIYCSKLTWKSRQADVDTCRMKGX- KDCBHFIFKVLKKNDDALFVCGTNA 139
Db 82 L-----VNKDFOKITWPSVSTRRDECKMAGKILKECANFIVKLVANVTHLYAGGTGA 136
Cc CC
Qy 140 FNPSCR-----NVMQDTLEPFGEDEF--SGARCPYDANKANVLFADGKLYSATVDFLA 192
Db 137 FHPICTYIEIGHNEPDNIFLENSHPENGSGSPYDKLLTASLLIDGELYSSTADEFMG 196
Qy 193 IDAVIYKSLGESPTLRITVKHDSKWLKEPYVOA-----VDYGDYIYFFPREIAVEYNT 245
Db 197 RDAIFRTLLGHHPIRTEGDSRLNDPKFISALHISEDNPPDDKYFFEFENALDGEH 256
Qy 246 MGKVPFRVAVOVCKNDMGSSQRYLEKQWTFKLARLNGSPG----DSHYFNILQAVTD 301
Db 257 SGRATHARIQOIKNDPFGG--HRSIVNMKTFLKARLIGSVGPNIGIDTFR-----DELD 310
Qy 302 VIRINGRD-----VVLATFSTPPYNSIPGSAYCAVMDLIDAVFTGRFKEQKPSDTWTPV 357
Db 311 VFLMNEKDPKPNVYVYVFTTSSNIFKGSAYCWSMSVRYKVLGPYAHKRGDPYQWVY- 369
Qy 358 DERVPKRPCCAGSSSLERYATSNPPDITLNFIKTHPLMDEAVPSIFNRPWFLRTWYR 417
Db 370 QGRVYPRPFCRP--SKTFGGFDSITKOLPDDVITPASHPAMVNPVFPNNRPVYIKTDVN 428
Qy 418 YRLTKIADVTAAGRYQNHVTFLESGEGLIKFLARIGNSGFLNDSLFLEMEVYNSK 477
Db 429 YQFQIYVDRVADAGQYDVWFITGTVGLKVASIPKETWYDLEEVLEEMTVFR---- 484
Qy 478 SYDVEDKRIKMGOLDRASSSLVYAFSTCVIKVPLRGCEHKGCKTICIASRDPYCGMIK 537
Db 485 -----EPFAISAMELSTKQOOLYIGSTAGVAGQPLMHCDIYGKACACCLARDPYCAW-- 537
Qy 538 EGAGCSHLSNPS-ELTEQDIERGNTDGLGDC---HNSFVALNGHS--SSLPSTTSD 590
Db 538 DGSACSRYFPFAKRTTRQDIRNG--DPLTFCSDLHHDNH---HGSPEERILYGVENSS 592
Qy 591 STAGEVYESRGMTDMKHLDSPPDTPPLGAVSSHNDKKGVRES 637
Db 593 TFLCSPKSRALVYQFORNERKEEI--RVDHIIIRTQGLLRS 638

RESULT 14
SKLA_TRICF STANDARD, PRT, 712 AA.
ID SKLA_TRICF Q26972;
AC 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Semaphorin 1A precursor (Semaphorin-1).
GN SEMA-1A OR TSEMA-1.
OS Tribolium confusum (Confused flour beetle).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
OC Tenebrionidae; Tribolium.
OC NCBI_Taxid=7071;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=94094332; PubMed=8269517;
RA Kolodkin A.L., Mathes D.T., Goodman C.S.;
RT "The semaphorin genes encode a family of transmembrane and secreted
RT growth cone guidance molecules."
RL Cell 75:1389-1399(1993).
CC -1- FUNCTION: PLAY A ROLE IN GROWTH CONES GUIDANCE.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -1- SIMILARITY: Contains 1 Sema domain.
Cc This SWISS-PROT entry is copyright. It is produced through a collaboration
Cc between the Swiss Institute of Bioinformatics and the EMBL outstation-
Cc the European Bioinformatics Institute. There are no restrictions on its

```

```

Cc use by non-profit institutions as long as its content is in no way
Cc modified and this statement is not removed. Usage by and for commercial
Cc entities requires a license agreement. (See http://www.isb-sib.ch/announce/
Cc or send an email to license@isb-sib.ch.)
Cc
Cc EMBL; L26080; AAA16609.1; -
Cc InterPro; IPRO03659; Plexin-like.
Cc InterPro; IPRO02165; Plexin_repeat.
Cc InterPro; IPRO01627; Sema.
Cc Pfam; PF01437; PSI; 1.
Cc SMART; SM00423; PSI; 1.
Cc SMART; SM00630; Sema; 1.
Cc Signal; Developmental protein; Transmembrane; Glycoprotein;
Cc Neurogenesis. 1
Cc SIGNAL 1 20
Cc CHAIN 21 712
Cc DOMAIN 21 601
Cc TRANSMEM 602 622
Cc DOMAIN 623 712
Cc DOMAIN 213 506
Cc DOMAIN 470 473
Cc DOMAIN 612 616
Cc CARBOHYD 42 42
Cc CARBOHYD 69 69
Cc CARBOHYD 161 161
Cc CARBOHYD 265 265
Cc SEQUENCE 712 AA; 79751 MW; C0734FE5B9C09FE3 CRC64;

Query Match 15.5%; Score 845; DB 1; Length 712;
Best local Similarity 28.9%; Pred. No. 8e-47;
Matches 230; Conservative 130; Mismatches 270; Indels 166; Gaps 30;

Qy 12 LHFAGAPDEPDESEPT-----SISGNYTKQYVFGHKRGNTGRHRLDIQWIMIMN- 65
Db 12 LIALCHAMPDSSSKLINHPKSVESKST-----GNATFPDH-----FYLNG 54
Qy 66 --GLTYIARADHIYTVDI--DTSHTTEIYCSKLTWKSRQADVDTCRMKGKDCCHNFIK 122
Db 55 DETSILVGGNRRYNNISIFPLSERK-----GGRIDWSPDAHGOLCLIKGTATDDCCQYIR 110
Qy 123 VLIKKNDDALFVCGTNAFNSGCRNYKMDTLEPPGD--FSGMARCPYDANKANVLPADGK 161
Db 111 ILYSSEPGKLVIGTNSYKPLCRTYAFKKGKYLVEKEVEIGICPPYNEHNSSTSVSNGQ 170
Qy 182 LVGATYTDPLAIDAVIYRS/GESPTLRITVHDSKWLKEPYVOAVDGYIYFFPREIAV 241
Db 171 LFSATVADSGGPPLYR-----PORTELSDIKQLNAPFVNSVAYGVDTIFFFYRETA 225
Qy 242 EYNTMGKVPFRVAVOVCKNDMGSSQRYLEKQWTFKLARLNGSPGDSHYFNILQAVTD 301
Db 226 EYNNCGKVIYSRAVAVCKDKGPGHGRDR--WTSFLKARLNGSIPGEYFPPEIOTSTD 284
Qy 302 VI--RINGRD-----VVLATFSTPPYNSIPGSAYCAVMDLIDAVTGRFKEQKPSDTWTPV 356
Db 285 IVEGRVNSDSSKTIYIGILTPVNAIGSALICAYQADILRVBGSFKHOETINSMLPV 344
Qy 357 PDERVPKRPCCAGSSSLERYATSNPPDITLNFIKTHPLMDEAVPSIFNRPWFLRTWY 416
Db 345 PQLNVPRPFGQCYRDSRI-----LPKQVNFIKTHLMED--VALGKPLVAVSL 395
Qy 417 RYRLTKIADVTAAGRYQNH--TVFLESGEGLIKFLARIGNSGFLNDSLFLEMSYVNS 474
Db 396 QYRFTALTVDPOVKTINNOYLDVLYIGTIDGKVLK----- 430
Qy 475 EKCSYDVEDKRIKMGOLDRASSSLVYAFSTCV-----IKVP----- 511
Db 431 -----ANITPRHAKALLYKRTISVHPHAPYKQLKIAGYGVVYVVGDEI 478
Qy 512 ---LGRCEHKGCKTICIASRDPYCGMIKEGAGCSHLSNPSRLTF--EQDIERGN----- 561
Db 479 RLALNLNHCASKTRC--KQCVLEQDPHCAMDAKQNLVCVIDTVTSYRFLIQDVVRGDDNKCW 537

```

QY 562 ----TDGLDCHNSFVALNGHSSSLPSTTSDSTAQEGYSGRMGLMDKLLDSPDTP 618
 Db 538 SPQTDKKTIVYK-----PSEVENET-----NSIDEXDL-----DSSDP 572
 QY 619 LGAVSSHNDKQKVIRESYLKG---HDQVPTV---LTAIVILFVWGA---VFSGITV 670
 Db 573 LKGTGDDSDSCPV-SENSIGCAVRCQGLVITAGTLHVVVVSIVGLFEMVLSGLSV 631
 QY 671 YCV--CDHRKDVAVVQKE--KELTHSRGNS-SVTLGSLFGDTQSKDPKPEALITP 725
 Db 632 FAKFHSDSQPAPFLEQNHLERLSANQGYLTPRANKAVNLVWVSSSTPRPKMD 691
 QY 726 LMHNGKATPNTAKM 741
 Db 692 VKKDNLIASDGLQKI 707

RESULT 15
 SZIB BRARE STANDARD; PRT: 778 AA.
 AC 09M66;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Semaphorin Z1B precursor (Semaphorin 1B) (Sema-Z1B).
 GN SEMAZIB OR SEMAZIB.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=99425174; PubMed=10495275;
 RA Roos W.; Schachner W.; Bernhardt R.R.;
 RT "Zebrafish semaphorin Z1b inhibits growing motor axons in vivo.";
 RL Mech. Dev. 87:103-117(1999).
 CC - FUNCTION: MIGHT NORMALLY INFLUENCE THE MIDSEGMENTAL PATHWAY CHOICE
 CC OF THE VENTRALLY EXTENDING MOTOR AXONS BY CONTRIBUTING TO A
 CC REPULSIVE DOMAIN IN THE POSTERIOR SOMITE.
 CC - SUBCELLULAR LOCATION: Secreted (by similarity).
 CC - TISSUE SPECIFICITY: EXPRESSED IN RHOMBOMERES THREE AND FIVE, AND
 CC IN THE POSTERIOR HALF OF NEWLY FORMED SOMITES WHICH IS AVOIDED BY
 CC VENTRALLY EXTENDING MOTOR AXONS.
 CC - SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC - SIMILARITY: Contains 1 Sema domain.
 CC - SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL: AF083382; AAD28103.1; -
 DR ZFIN: ZDB-GENE-991209-6; sema3ab.
 DR InterPro: IPR007110; IG-1-like.
 DR InterPro: IPR003599; IG.
 DR InterPro: IPR003066; IG_MHC.
 DR InterPro: IPR001659; Plexin-like.
 DR InterPro: IPR002165; Plexin_repeat.
 DR InterPro: IPR001627; Sema.
 DR Pfam: PF00047; Ig_1.
 DR Pfam: PF01437; PSI_1.
 DR Pfam: PF01403; Sema; 1.
 DR SMART: SM00409; PSI; 1.
 DR SMART: SM00423; PSI; 1.
 DR SMART: SM00630; Sema; 1.
 DR PROSITE: PS50835; IG_LIKE; 1.
 DR Signal: Immunoglobulin domain; Multigene family; Neurogenesis;
 KW Developmental protein; Glycoprotein.

FT SIGNAL 1 17 POTENTIAL.
 FT CHAIN 18 778 SEMAPHORIN Z1B.
 FT DOMAIN 241 539 SEMA.
 FT DOMAIN 579 668 IG-LIKE C2-TYPE.
 FT DOMAIN 721 776 ARG/LYS-RICH (BASIC).
 FT DISTLFD 652 716 BY SIMILARITY.
 FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 593 593 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 778 AA; 88904 MW; 4036FA323AE21895 CRC64;

Query Match 15.5%; Score 845; DB 1; Length 778;
 Best Local Similarity 34.3%; Pred. No. 9e-47;
 Matches 194; Conservative 94; Mismatches 204; Indels 74; Gaps 17;

QY 66 GTLYIARDHIYTV-DTSHTEIYCSKLTWKSQAQVDTCRMKGRH-KDECHNFIKV 123
 Db 66 GRLFVGAKDHLSEFLVDINNDDQL-----ISMPSPRRDECKKAGKDVCKECAFIV 122
 QY 124 LKKKNDALFVCGTNAFNPSCRNKYM-----DLEPPGDEF-SGMARCPYAKHANVALF 177
 Db 123 LQPFNQTHLYACGTGAFHPCAHAVEGKRSEDTFRLLGSSFENGKSPYDPKLOTASVL 182
 QY 178 ADGKLYSATVDELALDAIVIRSLGESPLTLTKHDSKWLKEPPYV-----QAVDY 228
 Db 183 IDGELYAGTSADFWRDPALFRLLGKHPDITREQDSRLNDPRFVSVHLIPESDPAED- 241
 QY 229 GDYIYFFREIAVEYNTGKVFPRVAVQVCKNDMGSGRVLEKQWTSFLKARLNSVPG- 287
 Db 242 -DKIYLFPRENALDDEQSKATYHARIGQLCKNDFGG-HRSLVKNKTTFLKARLVCSVPL 299
 QY 288 ---DSHFYFNILQAVTVDIRNGRD---VVLATSTYNSLPGSAVCAVMDLDAVPT 340
 Db 300 NGIDTHF-----DELQDVFLMSKDKPKNPITYAVFTSSNIFKGSAAVCMYSMADIRVFL 354
 QY 341 GRFKQKSPDSTWTPVDPDERVPKPRPGCCAGSSSLERATSNFEPDDTLNFIKTHPLME 400
 Db 355 GPYAHRODPNQWVPELN-RVPIYRPGTC-SKITDGESEIKDFDDVDVITPARSHIPAMYN 412
 QY 401 AVPSIENFPLRTWVRRLTKLIADVTAAGYQNHNTVFLGSEKGIILKFLARINSGL 460
 Db 413 PVPEPINNPILIKTDVDYQFOIVDREARDGQYDVAFIGTDMQTVLKVSIPRGTWHD 472
 QY 461 NDSLFLFEMSYNNEKSCYDGEVDEKRTMGMDLDAASSLVAFAFCVYKVPGRGERGK 520
 Db 473 LBEVLEEMTVFR-----EPTALTAMELSTKQGLYGSATGVSQMPHRCDDVYVK 523
 QY 521 CKTICIASRDPYCGMIKKGACSHSPNS-RLTFEODIERGNTDGLDCHNSFVALNGHS 579
 Db 524 ACAECCLARDPYCAM--DGSQCSRYFPYAKRRTTRQDIRND----- 563
 QY 580 SLLPSTTSDSTAQEGYSGRMGLMD 605
 Db 564 ----FLTQCSDLQHHDADGEAGLID 585

Search completed: October 23, 2003, 17:09:37
 Job time : 22 secs